சுட்டுறுப்பைச் சாராத முறைகள்

(NON - PARAMETRIC METHODS)

வி.எஸ். சம்பத்குமார்

தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

த. பா. நி. (க.வெ.) வரிசை எண்—726

சுட்டுறுப்பைச் சாராத முறைகள்

(பட்டப்படிப்பிற்குரியது)

வி. எஸ். சம்பத்குமார், எம்.எஸ்சி. (புள்ளியியல்),
துணைப்பேராசிரியர், புள்ளியியல் துறை,
மாநிலக் கல்லூரி,
சென்னை.

தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

First Edition—October, 1976

T.N.T.B.S. (C.P.) No, 726

© Government of Tamilnadu

NON—PARAMETRIC METHODS

V. S. SAMPATHKUMAR

Price Rs. 5-10

Published by the Tamilnadu Textbook Society under the Centrally Sponsored Scheme of Production of books and literature in regional languages at the University level, of the Government of India in the Ministry of Education and Social Welfare (Department of Culture), New Delhi.

Printed out of the Paper allotted by the Government of India.

Printed by
BHAGAT PRINTERS,
55, Railway Station Road,
Alandur, Madras-600 016.

பதிப்புரை

சுட்டுறுப்பைச் சாராத முறைகள் என்ற இந் நூல், தமிழ்நாட்டுப் பாடநூல் நிறுவனத்தின் 726ஆவது வெளியீடாகும். கல்லூரித் தமிழ்க் குழுவின் சார்பில் வெளியான 35 நூல்களையும் சேர்த்து இதுவரை 761 நூல்கள் வெளிவந்துள்ளன. இந் நூல் மைய அரசு, கல்வி, சமூக நல அமைச்சகத்தின் 'மாநில மொழியில் பல்கலைக்கழக நூல்கள் வெளியிடும் திட்டத்' தின்கீழ் வெளியிடப்படுகிறது.

மேலாண்மை இயக்குநர்
தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

பொருளடக்கம்

		பக்கம்
	தோற்றுவாய் (Introduction)	1
1.	ஒருகூறுச் சோதனைகள் (One Sample Tests) ...	10
2.	இரு தொடர்புள்ள மாதிரிகளுக்கான சோதனைகள் (Related Two Sample Tests)	44
3.	இரு சார்பற்ற கூறுகளுக்கான சோதனைகள் (Tests for Two Independent Samples)	80
4.	*k*-தொடர்புடைய கூறுகளுக்கான சோதனைகள் (Tests for *k*-related Samples)	150
5.	*k*-சார்பற்ற கூறுகளுக்கான சோதனைகள் (Tests for *k*-Independent Samples) ...	162
6.	ஒட்டுறவின் அளவைகளும் அவைகளுக்கான மிகைத் தன்மைச் சோதனைகளும் (Measures of Correlation and their Tests of Significance) ...	190
	பிற்சேர்க்கை (Appendix) ..	249
	மேற்கோள் நூற்பட்டியல் ...	254
	கலைச்சொற்கள்	257

தோற்றுவாய்

(Introduction)

நவீன புள்ளியியலின் மையத் தலைப்பாகப் புள்ளியியல் உய்த்துணர்வு திகழுகிறது. முழுமைத் தொகுதியின் சுட்டுறுப்பு களை மதிப்பிடுதல் (estimation of population parameters), எடு கோள் சோதனைகள் (tests of hypotheses) என்பன புள்ளியியல் உய்த்துணர்வின் இரு பகுதிகளாகும். பெருமளவிலுள்ள நிகழ்ச்சி களைப்பற்றி, அவைகளின் ஒரு பகுதியை மட்டும், கண்டறிதல் (observation) மூலமாக எவ்வாறு முடிவு கொள்வது என்பதைப் பற்றி அமைவதே புள்ளியியல் உய்த்துணர்வின் செயலாகும். இதற்குப் புள்ளியியலானது உற்ற கருவியாக, உருவாக்குதல் செய்து (formalize) முறைகளைத் திட்டப்படுத்தி முடிவுகளைக் கொள்வதற்கு உதவுகிறது. பொதுவாகப் புள்ளியியல் உய்த்துணர் வானது நிகழ்தகவு மூலமாக முழுமைத் தொகுதியிலிருந்து எடுக்கப் பட்ட மாதிரிக் கூறுகள் உண்மையிலேயே வேறுபட்டுள்ளனவா என்பதை இரு மாதிரிக் கூறுகளுக்கிடையேயுள்ள கண்டறிந்த வேறுபாடுகள் (observed differences) மிகைத்தன்மை செய்கிறதா (signify) என்பதைத் தீர்மானிக்கலாம். எப்பொழுதும் ராண்டம் முறையில் சேகரிக்கப்படும் இரு குழுவின் மதிப்புகள் ஓரளவுக்கு வேறுபட்டிருப்பதைக் காண முடியும். இவ்வித வேறுபாடுகள் தற்செயலாக ஏற்படக் கூடியவைகளாக இருக்கலாம். கொடுக்கப் பட்ட எவ்விதப்பட்ட வகைக்கும், ஏற்படக்கூடிய கண்டறிந்த வேறுபாடுகள் தற்செயலாக ஏற்பட்டவையா, இல்லையா என் பதைப் புள்ளியியல் உய்த்துணர்வின் வாயிலாக மாதிரிக் கூறுகளின் முழுமைத் தொகுதியின் தன்மைக்கு ஏற்ப தற்கோளை (assumption) அமைத்து முடிவுகளைக் காண்கிறோம். முழுமைத் தொகுதியின் மதிப்புகள் 'சுட்டுறுப்புகள்' (parameters) என்பதால் இந்தப் புள்ளியியல் முறைகளைச் சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் முறைகள் என்கிறோம். உதாரணமாக, கூறு மாதிரிகளின் மதிப்புகள் இயல்நிலைப் பரவலைப் பெற்ற முழுமைத் தொகுதியிலிருந்து வரப் பெற்றவை என்ற தற்கோளைக் கருதுவதன்மூலம் உய்த்துணரலாம். சமீப காலத்தில் உய்த்துணர்வின் மிகப் பெருமளவிலான முறை களில் சுட்டுறுப்புகளைப்பற்றின தற்கோள்கள் கடினமாக இல்லா

திருப்பதைக் காணலாம். இம்மாதிரியான சுட்டுறுப்புகளைச் சாராத முறைகள் அல்லது பரவல் விடுபட்ட முறைகள் (distribution free methods) எளிதாக முடிவுகளைக் கொள்வதற்கு வழி வகுக்கின்றன. சில சுட்டுறுப்புச் சாராத முறைகளை எப்பொழுதும் 'தரச் சோதனைகள்' (ranking tests) அல்லது 'வரிசைச் சோதனைகள்' (order tests) என்று கூறுகிறோம்.

தீர்மானத்திற்கான நோக்க முறைகள் (Objective procedures for decision)

ஒரு கட்டான விவரங்களுக்கு ஒரு குறிப்பிட்ட எடுகோளை உறுதி செய்வதற்குத் தீர்மானத்திற்கான நோக்க முறைகளைக் கொள்வது மிகவும் அவசியமாகிறது. அதாவது ஓர் எடுகோளை ஏற்றுக்கொள்ளவோ புறக்கணிக்கவோ தீர்மானத்திற்கான நோக்கமுறை இன்றியமையாததாகிறது. அவைகளாவன: (a) சூனிய எடுகோளான H_0-ஐ அமைப்பது, (b) H_0-ஐச் சோதனை செய்வதற்குப் புள்ளியியல் சோதனையை [புள்ளியியல் மாதிரி வடிவத்திற்கு ஏற்ற (statistical model) வகையில்] தேர்ந்தெடுப்பது, (c) α என்ற மிகைத்தன்மை மட்டத்தையும் (level of significance) மாதிரி அளவான Nஐயும் தெரிவிப்பது, (d) H_0-ன் அடிப்படையில் கொள்ளப்பட்ட புள்ளியியல் சோதனையின் மாதிரிப் பரவலைக் (sampling distribution) காண்பது, (e) பிறகு (*b*), (*c*), (*d*) ஆகியவைகளுக்கேற்றதாக அமையக்கூடிய மறுக்கத்தகும் பகுதியை (rejection regions) வரையறை செய்வது, (f) கொடுக்கப்பட்ட விவரங்களிலிருந்து புள்ளியியல் சோதனை மதிப்பைக் கண்டு, அம் மதிப்பு மறுக்கத்தகும் பகுதியில் அமைந்திருந்தால் H_0-ஐ அம் மிகைத்தன்மை மட்டத்தில் மறுப்பது போன்றவைகளாகும்.

முழுமைத் தொகுதியின் புள்ளியியல் அளவையின் (parameter) மதிப்புத் தெரியாதபோது அது ஒரு குறிப்பிட்ட நிலையெண்ணாக இருக்கலாம் என்று ஊகம் செய்து அம் முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட கூறு விவரங்களைக் கொண்டு அவ்வூகம் ஏற்றுக்கொள்ளத் தக்கதுதானா அல்லது மறுக்கத்தக்கதுதானா எனச் சோதனை செய்து முடிவுகொள்ள ஏதுவான வழி முறைகளை எடுகோள் சோதனைகள் என்று கூறுகிறோம்.

கூறு உறுப்புகள் அவைகளின் முழுமைத் தொகுதியின் பண்புகளை உணர்த்துமாறு அமைகின்றன. எனவே, கூறின் அடிப்படையில் பெற்ற புள்ளியியல் அளவையின் மதிப்பானது அதை ஒத்த முழுமைத் தொகுதி அளவையின் மதிப்பை அணுகி அமையும் என்று பொதுவாக எதிர்பார்க்கிறோம். ஆனால், கூறில் சிற்சில

உள்மாற்றங்களினால் புள்ளியியல் அளவையின் மதிப்பு முழுமைத் தொகுதியின் அளவையிலிருந்து வேறுபடுகிறது. பொதுவாக, இத்தகைய வேறுபாடுகள் மறுக்கத்தக்க வகையில் அமைந்துள்ளன. அதனால் புள்ளியியல் அளவையின் மதிப்பு முழுமைத் தொகுதி அளவையின் மதிப்பை அணுகி அமைவதற்கு அதிகமான வாய்ப்பையும், மிகுந்து விலகியிருப்பதற்குக் குறைந்த வாய்ப்பையும் பெற்று அமைகிறது என்று கூறுகிறோம்.

இனி, முழுமைத் தொகுதியின் அளவைக்கு ஓர் ஊகமான மதிப்பை அளித்து, அவ்வூகம் சரிதானா என்பதைச் சோதனை செய்ய விழைகிறோம். அம் முழுமைத் தொகுதியிலிருந்து பெறப்பட்ட கூறின் அளவையின் மதிப்பு இவ்வூக மதிப்பிற்கு அணுகியிருந்தால் இவ்விரு மதிப்புகளுக்குமிடையேயுள்ள வேறுபாட்டைக் கூறின் சிற்சில உள்மாற்றங்கள் ஏற்படுத்தியதாகக் கொண்டு ஊக மதிப்பை முழுமைத் தொகுதி அளவையின் சரியான மதிப்பெனத் தீர்மானிக்கிறோம். அவ்வாறின்றி, ஊக மதிப்பிற்கும் கூறின் அளவை மதிப்பிற்கும் இடையேயுள்ள வேறுபாடு மிகுதியாக இருக்கும் போது ஊக மதிப்பு முழுமைத் தொகுதி அளவையின் மதிப்பாக அமையவில்லை என்று தீர்மானிக்கிறோம். இம்மாதிரியான சோதனைகள் புள்ளியியல் எடுகோள் சோதனைகள் என்று கூறுகிறோம்.

முழுமைத் தொகுதியின் கூட்டுச் சராசரியை μ எனக் குறிப்பிட்டு, μ-ன் மதிப்புத் தெரியாதபொழுது அதனை μ_0 என ஊகமாகக் கொண்டு, கருதப்பட்ட ஊகம் சரியானதா என அம் முழுமைத் தொகுதியிலிருந்து பெறப்பட்ட கூறு வாயிலாகச் சோதனை செய்ய முற்படுகிறோம். இச் சோதனையின்போது $\mu = \mu_0$ என்பது சோதனைக்கான எடுகோளாகும். இவ்வாறாக, சோதனைக்காக அமைத்துக் கொள்ளும் எடுகோளை மறுக்கத்தகு எடுகோள் அல்லது சூனிய எடுகோள் (null hypothesis) என்று கூறுகிறோம். இரு கூறுகளின் சராசரிகள் முறையே μ_1, μ_2 எனில், $\mu_1 = \mu_2$ என்ற சூனிய எடுகோளைக் கருதினால், $\mu_1 \neq \mu_2$ என்ற மாற்று எடுகோளையோ $\mu_1 > \mu_2$, $\mu_1 < \mu_2$ என்ற மாற்று எடுகோள்களையோ கருதிச் சோதனை செய்யலாம். இந்த மாற்று எடுகோள்களில் $\mu_1 \neq \mu_2$ என்ற எடுகோள் இருமுனைச் சோதனையையும் $\mu_1 > \mu_2$, $\mu_1 < \mu_2$ என்ற எடுகோள்கள் ஒருமுனைச் சோதனையையும் குறிக்கும்.

சூனிய எடுகோளையும், மாற்று எடுகோளையும் அமைத்த பிறகு புள்ளியியல் சோதனையை அப் புள்ளியியல் மாதிரி வடிவத்திற்கு ஏற்றவாறு தேர்ந்தெடுக்கிறோம்.

இப் புள்ளியியல் சோதனையை மேற்கொள்ளும்பொழுது முழுமைத் தொகுதியின் அளவைகளைப்பற்றி எந்தவிதமான தற்கோளையும் (assumption) நாம் மேற்கொள்ளுவதில்லை. எனவே, இம்மாதிரியான புள்ளியியல் சோதனை சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனையாக அமைகிறது.

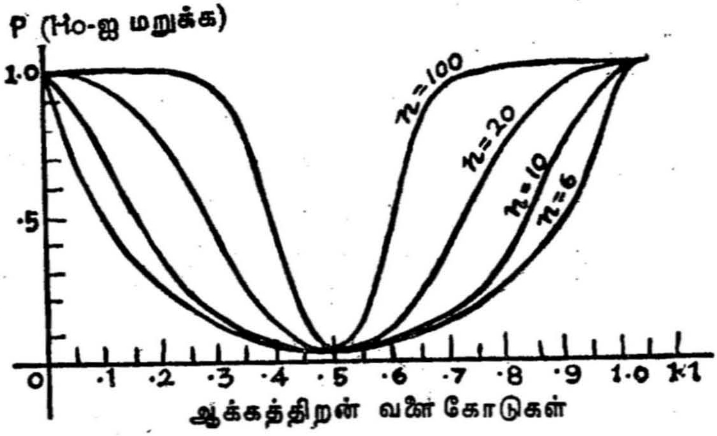
H_0 என்ற சூனிய எடுகோளும், H_1 என்ற மாற்று எடுகோளும் அறிவிக்கப்பட்ட பிறகு α என்ற மிகைத்தன்மை மட்டமும் N என்ற மாதிரிக் கூறின் அளவும் தெரிவிக்கப்பட வேண்டும். விவரங்கள் சேகரிக்கப்படும் முன்பே H_0 உண்மையாக இருக்கும் பொழுது ஏற்படக்கூடிய முடியுமான மாதிரிகள் அனைத்தையும் தெரிவித்துவிடுகிறோம். இவைகளிலிருந்து நிகழ்தகவு மிகவும் சிறியதாயிருக்கும் வண்ணம் புறக்கோடியாக (extreme) அமைந்துள்ள இயலுமான உபமாதிரிகளைத் தெரிவிக்கிறோம். மேலும் H_0 உண்மையாக இருக்கும்போது நாம் கண்டறியும் மாதிரியானது இவைகளில் ஒன்றாக இருக்கும். நம்முடைய புள்ளியியல் சோதனை அளிக்கும் மதிப்பானது H_0-ன் கீழ் நிகழக்கூடிய நிகழ்ச்சியின் நிகழ்தகவின் மதிப்பான α-க்குச் சமமாகவோ குறைவாகவோ இருந்தால் H_0-ஐ மறுத்து H_1-ஐ ஏற்றுக்கொள்கிறோம். இந்தச் சிறிய நிகழ்தகவு மதிப்பான α-ஐ மிகைத்தன்மை மட்டம் (level of significance) என்று கூறுகிறோம்.

H_0-ஐப்பற்றி முடிவு கொள்ளும்போது இரண்டுவிதமான பிழைகள் நேரக்கூடும். இதில் H_0 உண்மையாக இருக்கும் பொழுது அதை மறுப்பது என்பது முதல்வகைப் பிழையாகும். இதனை வகை I பிழை (Type-I error) என்று கூறுகிறோம். H_0 பொய்யானதாக இருக்கும்போது அதை ஏற்றுக்கொள்ளுவது என்பது இரண்டாம் வகைப் பிழையாகும். இதனை வகை II பிழை (Type-II error) என்று கூறுகிறோம். முதல் வகைப் பிழை ஏற்படுவதற்கான நிகழ்தகவின் மதிப்பை α என்றும் இரண்டாம் வகைப் பிழை ஏற்படுவதற்கான நிகழ்தகவின் மதிப்பை β என்றும் குறிப்பிடுகிறோம். α-ம், மாதிரியின் அளவான N-ம் தெரிந்தால் β-ஐக் காண முடியும். இவ்விரு பிழைகளையும் குறைத்து N-ஐ அதிகரிக்க வேண்டும்.

H_0 பொய்யாக இருக்கும்பொழுது அதனை மறுப்பதற்கான நிகழ்தகவைச் சோதனையின் ஆக்கத்திறன் (power of a test) என்று அழைக்கிறோம். அதாவது

ஆக்கத்திறன் = 1— (இரண்டாம் வகைப் பிழை நிகழ்வதற்கான நிகழ்தகவு)
= 1—β ஆகும்.

கீழுள்ள படத்தில் $N=6, 10, 20, 100$ ஆகியவைகளுக்கான ஆக்கத்திறன் வளைகோடுகள் அமைக்கப்பட்டுள்ளன. இதில் N-ன் அளவு அதிகமாக ஆகச் சோதனையின் ஆக்கத் திறனும் அதிக மாவதைக் காணலாம்.

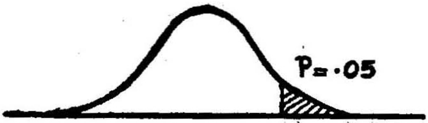


ஆக்கத்திறன் வளை கோடுகள்

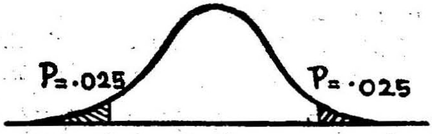
இப்பொழுது H_0-ன் கீழ் அமைந்துள்ள புள்ளியியல் சோதனைக் கான மாதிரிப் பரவலைக் காணவேண்டும். மாதிரிப் பரவல் என்பது ஓர் அறிமுறைப் பரவல் (theoretical distribution) ஆகும். அதாவது, H_0-ன்கீழ் ராண்டம் முறையில் எடுக்கப்பட்ட சமமான கூறு அளவைப் பெற்ற மாதிரிகளிலிருந்து கணக்கிடப்பட்ட ஏதோ ஒரு புள்ளியியல் அளவையின் (உதாரணமாக $\bar{X}$) இயலுமான மதிப்புகளின் பரவலாகும். H_0-ன் கீழ் புள்ளியியல் அளவையின் பல்வேறு இயலுமான எண்ணளவியலான மதிப்புகளோடு தொடர் புடைய நிகழ்தகவு மதிப்புகளைப் புள்ளியியல் அளவையின் மாதிரிப் பரவல் குறிக்கிறது. இந்த மாதிரிப் பரவலின் அடிப்படையில் H_0-ஐ ஏற்றுக்கொள்ளவோ மறுக்கவோ செய்கிறோம். மறுக்கத்தகும் வெளியே (region of rejection) மாதிரிப் பரவலின் வெளியாகும். மாதிரிப் பரவலானது H_0-ன் கீழ் சோதனைப் புள்ளியியல் அளவையின் (test statistic) இயலுமான எல்லா மதிப்புகளையும் உள்ளடக்கி இருக்கும். மறுக்கத்தகும் வெளியானது இந்த இயலுமான மதிப்புகளின் உப கணங்களைப் (sub-sets) பெற்று அமையும். இந்த உபகணத்தில் சோதனைப் புள்ளியியல் அளவை, H_0-ன் கீழ் அமைவதற்கான நிகழ்தகவை α என்று குறிக்கிறோம். எனவே, மறுக்கத்தகும் வெளியில் அமைந்துள்ள மதிப்போடு தொடர்பு கொண்டுள்ள நிகழ்தகவானது α-க்குச் சமமாகவோ குறைவாகவோ இருக்கும்.

மறுக்கத்தகும் வெளியின் இருப்பிடம் (location) மாற்று எடுகோளான H_1-ன் தன்மையைப் பொறுத்து அமைகிறது. H_1 ஆனது ஒருமுனைச் சோதனையையோ இருமுனைச் சோதனையையோ குறிப்பிட்டால் மறுக்கத்தகும் வெளியானது ஒரு முனைச் சோதனைக்கும் இருமுனைச் சோதனைக்கும் மாறுபட்டு அமைகிறது. ஒருமுனைச் சோதனைக்கு மறுக்கத்தகும் வெளியானது மாதிரிப் பரவலின் ஒரு முனையில் அமைந்திருக்கும். இருமுனைச் சோதனையாக இருந்தால் மறுக்கத்தகும் வெளியானது மாதிரிப் பரவலின் இரு முனையிலும் அமைந்திருக்கும்.

மறுக்கத்தகும் வெளியின் அளவு $\alpha = 0 \cdot 05$ எனில் மறுக்கத் தகும் வெளியின் அளவானது மாதிரிப் பரவலின் கீழ் அமைந்துள்ள பரப்பில் 5 சதவீதம் ஆகும். $\alpha = 0 \cdot 05$-க்கு ஒருமுனைக்கான மறுக்கத்தகும் வெளியும், இருமுனைக்கான மறுக்கத்தகும் வெளியும் கீழுள்ள படத்தில் கொடுக்கப்பட்டுள்ளன.



$\alpha = 0.05$க்கு ஒரு முனைக்கான மறுக்கத்தகும் வெளி



$\alpha = 0.05$க்கு இரு முனைக்கான மறுக்கத்தகும் வெளி

புள்ளியியல் அளவையின் சோதனை அளிக்கும் மதிப்பானது மறுக்கத்தகும் வெளியில் அமைந்திருந்தால் H_0-ஐ மறுக்கிறோம். அதாவது புள்ளியியல் சோதனையின் கண்டறிந்த மதிப்பிற்குத் தொடர்புடைய நிகழ்தகவின் மதிப்பானது முன்பே ஊகித்துக் கொண்ட α என்ற மட்டத்திற்குச் சமமாகவோ குறைவாகவோ இருந்தால் H_0 பொய்யானது என்று முடிவு கொள்கிறோம். அம்மாதிரி கண்டறிந்த மதிப்பானது மிகைத்தன்மையைப் (significant) பெற்றதாகக் கூறுகிறோம்.

சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனையைப் புள்ளியியல் மாதிரி வடிவத்திற்கு (statistical model) ஏற்றவாறு தேர்ந்தெடுக்கும்பொழுது, அவைகளில் பயன்படுத்தப்படும் மாறிகளைப் பல்வேறு

அளவுத்திட்ட அளவுகளாகப் (scales of measurement) பயன்படுத்த வேண்டியுள்ளது. அளவுகள் பல வகைகளாகவும், வேறுபட்ட அளவிற்குத் திருத்தமாகக் கொள்ளப்பட்டதாகவும் இருக்கலாம். ஒரு குழுவிலுள்ளவர்களை அவர்களின் உயரத்திற்கேற்ப வரிசையாக நிற்கச் செய்து, பார்த்த அளவிலேயே உயரத்தை மதிப்பிடலாம். இம்முறையின் மூலம் உயரத்தை மதிப்பிட்டால் *X* என்பவர், *Y* என்பவரைவிட உயரமானவர் அல்லது குட்டையானவர் என்று மட்டுமே கூறமுடியும். அதனுல் ஒவ்வொரு நபரின் உயரத்தையும் சரியாக அறிந்துகொள்ள முடியாது. ஓர் அளவுகோலின் மூலம் குழுவிலுள்ளவர்களின் உயரத்தை அளந்தால் ஒவ்வொரு நபரின் உயரத்தையும் தனித்தனியாக, துல்லியமாக அறிய முடியும். இம்முறையினுல் *X* என்பவரின் உயரம் *Y* என்பவரின் உயரத்தை விட எத்தனை மீட்டர் அல்லது சென்டிமீட்டர் அதிகமாக அல்லது குறைவாக இருக்கும் என்பதைத் திட்டவட்டமாகக் கூற இயலும். ஒரு பண்பை அளக்கும் செய்கையினுல் கண்டறிந்தவற்றின் அடிப்படையில் அளவுத் திட்டத்தை நான்கு வகைகளாகப் பிரிக்கலாம். அவைகளாவன: (i) பெயரளவான அளவுத் திட்டம் (Nominal scale), (ii) வரிசை அல்லது தர அளவுத் திட்டம் (Ordinal or Ranking scale), (iii) இடைவெளி அளவுத் திட்டம் (Interval scale), (iv) விகித அளவுத் திட்டம் (Ratio scale).

ஒரு குழுவிலுள்ள நபர்களை அவர்களின் உடலின் நிறத்திற்கு ஏற்பப் பிரித்தால், ஒருவருடைய உடலின் நிறம் மற்றொருவருடைய உடலின் நிறத்தைப் போலவோ வேறுபட்டோ இருக்கும் என்று மட்டுமே கூற இயலும். ஆனுல், உடலின் நிறத்தைக் கொண்டு அந் நபர்களை வரிசைப்படுத்தவோ எந்த அளவிற்கு அதிகமாக அல்லது குறைவாக இருக்கும் என்று கூறவோ முடியாது. எனவே, இங்கு உடலின் நிறம் பெயரளவான மாறியைக் (nominal variable) குறிக்கிறது. இக் குழுவிலுள்ள நபர்களை அடையாளம் காணக் குறியீடுகளையோ அல்லது எண்களையோ பயன்படுத்தினுல் அத்தகைய குறியீடுகள் அல்லது எண்கள் பெயரளவான அளவுத் திட்டத்தைக் குறிக்கும்.

ஒரு குழுவிலுள்ள நபர்களை அல்லது பொருள்களை ஒரு பண்பின் அடிப்படையில் தரப்படுத்தும்பொழுது (ranked) அல்லது 1, 2, 3,.........என்று வரிசையாக எண்களையிட்டு அமைக்கும் பொழுது அத் தர மதிப்பீடுகள் அல்லது வரிசை எண்கள், வரிசை அளவுத் திட்டத்தில் (ordinal scale) அமைகின்றன.

வரிசை அளவுத் திட்டத்திற்கான எல்லாத் தன்மைகளையும் பெற்றிருந்து, மதிப்புகள் சம அலகுகளில் (equal units) அமைக்கப்

பட்டிருக்கும்பொழுது அம் மதிப்புகள் இடைவெளி அளவுத் திட்டத்தில் (interval scale) அமைந்துள்ளனவாகக் கூறுகிறோம். இடைவெளி அளவுத் திட்டங்கள் உண்மையான பூச்சியப் புள்ளியை (true zero point) பெற்றிரா. ஒரு தேர்வில் ஒருவர் 0 என்ற மதிப்பைப் பெற்றால், அவருக்குத் திறமையே இல்லை என்பதை அம் மதிப்புக் குறிக்காது. எனவே, இடைவெளி அளவுத் திட்டத்தில் உண்மையான பூச்சியப் புள்ளியையும், அளவை அலகையும் (unit of measurement) யாதாமொரு (arbitrarily) முறையில் எடுக்கலாம்.

இடைவெளி அளவுத் திட்டத்திற்கான எல்லாத் தன்மைகளையும் பெற்றிருந்து, உண்மையான பூச்சியப் புள்ளியிலிருந்து அளவுகள் எடுக்கப்பட்டிருக்கும்போது, அந்த அளவுகள் விகித அளவுத் திட்டத்தில் (ratio scale) அமைகின்றன.

நாம் காணவிழையும் சுட்டுறுப்பைச் சாராத புள்ளியியல் முறைகளில் பயன்படுத்தப்படும் விவரங்கள், மேற்சொன்ன நான்கு அளவுத் திட்டங்களின் அடிப்படையில் அமைந்திருக்கும். அந்தந்த அளவுத் திட்டங்களுக்கு ஏற்பப் பல்வேறு சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனைகளைச் செய்து எடுகோள்களைச் சோதிக்கிறோம்.

மாதிரிகள் தேர்ந்தெடுக்கப்படும் முழுமைத் தொகுதியின் சுட்டுறுப்புகளுக்குச் சில நிபந்தனைகளைத் தெரிவிக்கும் மாதிரி வடிவத்தின் (model) சோதனையே, சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் சோதனையாகும் (Parametric Statistical Test). இதில் கருதப்படும் அளவைகள் யாவும் குறைந்தது இடைவெளி அளவுத் திட்டத்தில் இருக்கவேண்டும்.

மாதிரிகள் தேர்ந்தெடுக்கப்படும் முழுமைத் தொகுதியின் சுட்டுறுப்புகளுக்கு நிபந்தனைகள் ஏதும் தெரிவிக்காத மாதிரி வடிவத்தின் சோதனையே சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனையாகும் (Non-Parametric Statistical Test). ஆனால், சில தற்கோள்களைப் (assumptions) பெரும்பாலான சுட்டுறுப்பைச் சாராத சோதனைகளில் பயன்படுத்துகிறோம். உதாரணமாக, கண்டறிந்தவை சார்பற்றவை என்றும் ஆய்வில் கருதப்படும் மாறிகள் தொடர்ச்சியானவை என்றும் கொள்கிறோம். சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் சோதனைகளுக்குக் கருதப்படும் தற்கோள்கள் சுட்டுறுப்பைச் சாராத சோதனைகளுக்குக் கருதப்படும் தற்கோள்களைவிடப் பலவீனமானவைகளாகும். மேலும் சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் சோதனைகளைப் போலன்றி, சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனைகளில் பெயரளவான

அளவுத் திட்டம், வரிசை அல்லது தர அளவுத் திட்டம், இடைவெளி அளவுத் திட்டம் என்று பலதரப்பட்ட அளவுத் திட்டங்களைப் பயன்படுத்துகிறோம். பொதுவாகச் சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் சோதனைகளில், மாதிரி வடிவத்தில் கருதப்படும் தற்கோள்களின் நிபந்தனைகள் பூர்த்தி செய்யப்பட்ட மாறிகளின் அளவுத் திட்டம், இடைவெளி அளவுத் திட்டமாக அமைந்திருந்தால், அச் சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் சோதனையானது சக்தி வாய்ந்த சோதனையாக அமைகிறது. எனினும், சுட்டுறுப்பைச் சார்ந்த சோதனைகளுக்குத் தேவையான நிபந்தனைகள் யாவும் பூர்த்தி செய்யப்பட்டிருந்தாலும் மாதிரி அளவான N-ஐ ஒரு குறிப்பிட்ட அளவு அதிகப்படுத்திச் சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனைகளைச் செய்து அதே ஆக்கத்திறனைப் (power) பெறலாம்.

பெரும்பாலான சுட்டுறுப்பைச் சாராத புள்ளியியல் முறைகளில், கொடுக்கப்பட்ட விவரங்களைத் தரங்களாகவும் (ranks) குறிகளிட்டும் (signed) அமைத்துச் சோதனை செய்கிறோம். மாதிரியில் உள்ள தகவல்களைச் சுட்டுறுப்பைச் சார்ந்த முறைகள் மிகச் சரியான முறையில் பயன்படுத்துகின்றனவா அல்லது சுட்டுறுப்பைச் சாராத முறைகள் பயன்படுத்துகின்றனவா என்பதை அறிய அளவுகள் (measurements) அடிப்படையான காரணமாக அமைகின்றன. இடைவெளி அளவுத் திட்டத்தைவிட பலவீனமாக இருந்தால் கொடுக்கப்பட்ட விவரங்களின் கண்டறிந்தவைகளுக்குத் தரங்கள் அல்லது குறிகளை அளித்துச் சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனை செய்கிறோம். சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனைகளில், முழுமைத் தொகுதியின் பரவல் வடிவம் எப்படி இருப்பினும் சிறிய கூறுகளுக்கான நிகழ்தகவு மதிப்புகள் மிகத் துல்லியமான மதிப்புகளாக இருக்கும். ஆனால், பெரிய கூறுகளுக்குத் தோராயமான நிகழ்தகவு மதிப்புகளைப் பெறலாம். கூறின் அளவு 6-ம், அதற்குக் குறைவாகவும் இருந்தால் சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனையைப் பயன்படுத்த வேண்டிய நிலை ஏற்படுகிறது. மேலும் கொடுக்கப்பட்ட விவரங்கள் வரிசைப்படுத்தப்பட்ட நிலையில் இருந்தாலும் பெயரளவிலான அளவுத் திட்டத்தைப் பெற்றிருந்தாலும் சுட்டுறுப்பைச் சாராத புள்ளியியல் சோதனையைத்தான் பயன்படுத்தமுடியும் என்ற நிலை ஏற்படுகிறது.

கொடுக்கப்பட்ட விவரங்கள், ஒரு கூறாகவோ, இரு கூறாகவோ அல்லது பல கூறுகளாகவோ இருந்தால், சுட்டுறுப்புச் சோதனையை மேற்கொள்ளும் பல்வேறு சோதனைகளைப்பற்றிப் பின்வரும் அத்தியாயங்களில் காண்போம்.

1. ஒருகூறுச் சோதனைகள்

(One Sample Tests)

ஈருறுப்புச் சோதனை

ஈருறுப்புச் சோதனையைப் புள்ளியியல் ஆய்வுக்கேற்ற விவரங்களுக்கும் பயன்படுத்தலாம். சில சமயங்களில் ஈருறுப்புச் சோதனை மிகச் சக்தி வாய்ந்த சோதனையாகவும் திகழுவதுண்டு. மற்றத் தருணங்களில் சக்திவாய்ந்த சோதனைகள் இருப்பினும், இச் சோதனை சுட்டுறுப்பைச் சாராத புள்ளியியலுக்கு மிகவும் ஏற்ற சோதனையாகக் கருதப்படுவதுண்டு. மேலும், இச் சோதனை மிகவும் எளிதாக இருப்பதாலும், மறுக்கப்படவேண்டிய சூனிய எடுகோள் மறுக்கப்படுமளவிற்குச் சக்தி வாய்ந்த சோதனை என்பதாலும் இதனைப் பெரிதும் நாடுகிறோம். இப்போது இச் சோதனையைப் பற்றிக் காண்போம்.

மாதிரியானது (sample) ஒன்றையொன்று சார்பற்ற முயற்சிகளின் (independent trials) வெளிப்பாடுகளைப் (out comes) பெற்றதென்றும், ஒவ்வொரு வெளிப்பாடும் வகுப்பு 1-ஐ, p என்ற நிகழ்தகவுடனோ அல்லது வகுப்பு 2-ஐ, $(1-p)$ என்ற நிகழ்தகவுடனோ குறிக்கக் கூடியது என்றும் கொள்க.

வகுப்பு 1-ல் உள்ள கண்டறிந்தவைகளின் (observations) எண்ணிக்கை n_1 என்றும், வகுப்பு 2-ல் உள்ள கண்டறிந்தவைகளின் எண்ணிக்கை $n_2 = n - n_1$ என்றும் கொள்க. p^* ஆனது $0 \leqslant p^* \leqslant 1$ என்று இருக்குமாறு அமைந்துள்ள ஏதோ ஒரு குறிப்பிட்ட நிலையெண் எனில், எடுகோளானது கீழ்க்கண்ட மூன்று வடிவங்களைப் பெற்று அமையும்.

1. இருமுனைச் சோதனை (Two tailed test)

   $H_0: \quad p = p^*$

   $H_1: \quad p \neq p^*$

2. ஒருமுனைச் சோதனை (One tailed test)

$H_0 : p \leqslant p^*$

$H_1 : p > p^*$

3. ஒருமுனைச் சோதனை (One tailed test)

$H_0 : p \geqslant p^*$

$H_1 : p < p^*$

வகுப்பு 1-ன் வெளிப்பாட்டின் நிகழ்தகவைப்பற்றி மட்டும் கருதி அதற்கேற்ற சோதனைப் புள்ளியியல் அளவையை T எனக் கொள்க.

அதாவது, $T = n_1$ (1)

ஈருறுப்புப் பரவலின் இரு முனைக்கான தீர்வுகட்டமான வெளியின் அளவை (critical region of size) α-க்கு ஏற்றதான கீழ் முனையின் அளவு α_1, மேல் முனையின் அளவு α_2 ஆகும். அதாவது $\alpha_1 + \alpha_2 = \alpha$. எனவே, அட்டவணை D-லிருந்து குறிப்பிட்ட p^*, n ஆகியவைகளுக்கு ஏற்ற எண்ணுனை t_1-ஐ $P\,(Y \leqslant t_1) = \alpha_1$ என்றிருக்குமாறும், t_2-ஐ $P\,(Y \leqslant t_2) = 1 - \alpha_2$ என்றிருக்குமாறும் காண்கிறோம். இதில் Y ஆனது, p^*, n ஆகிய சுட்டுறுப்புகளைப் பெற்ற ஈருறுப்பு ராண்ட மாறியாகும். இதில் α_1-ம், α_2-ம் ஒன்றுக்கொன்று ஏறக்குறைய சமமாக இருக்க வேண்டும்.

T ஆனது t_2-க்கு அதிகமாக இருந்தாலோ அல்லது t_1-க்குக் குறைவாக அல்லது சமமாக இருந்தாலோ H_0-ஐ மறுக்கிறோம். இதற்கு மாறாக இருந்தால் H_0-ஐ ஏற்றுக் கொள்கிறோம். (2)-லுள்ள எடுகோளுக்கான ஒருமுனைச் சோதனைக்கு T-ன் உயர் மதிப்புகளுக்கு H_0 மறுக்கத்தக்கது என்பதால், α-ஐ அளவாகக் கொண்ட தீர்வுகட்டமான வெளியில் $T > t$ என்ற எல்லா T மதிப்புகளும் அடங்கி இருக்கும். இதில் t என்பது p^*, n ஆகிய சுட்டுறுப்புகளுக்கு ஏற்ற அட்டவணை மதிப்பாகும். இதனை அட்டவணை D-யிலிருந்து, $P\,(Y > t) = \alpha$ அல்லது $P\,(Y \leqslant t) = 1 - \alpha$ என்றிருக்குமாறு பெறலாம். இங்கு Y என்பது p^*, n ஆகிய சுட்டுறுப்புகளைப் பெற்ற ஈருறுப்புப் பரவலாக அமைகிறது. T-ன் மதிப்பு t-ஐ விட அதிகமாக இருந்தால் H_0-ஐ மறுக்கிறோம். T-ன் மதிப்பு t-ஐ விடக் குறைவாகவோ சமமாகவோ இருந்தால் H_0-ஐ ஏற்றுக் கொள்கிறோம்.

(3)-லுள்ள எடுகோளுக்கான ஒருமுனைச் சோதனைக்கு T-ன் சிறிய மதிப்புகளுக்கு, H_0 மறுக்கத்தக்கது என்பதால் α-யை அளவையாகக் கொண்ட தீர்வுகட்டமான வெளி $T \leqslant t$ என்ற

எல்லா T மதிப்புகளையும் பெற்றிருக்கும். இதில் t என்பது P^*, n ஆகிய சுட்டுறுப்புகளுக்கு ஏற்ற அட்டவணை மதிப்பாகும். இதனை அட்டவணை D-யிலிருந்து, $P(Y \leqslant t) = \alpha$ என்றிருக்குமாறு பெறலாம். இங்கு Y ஆனது p^*, n ஆகிய ஈருறுப்புப் பரவலுக்கான சுட்டுறுப்புகளைப் பெற்றுள்ளது. T-ன் மதிப்பு t-க்குக் குறைவாகவோ சமமாகவோ இருந்தால் H_0-ஐ மறுக்கிறோம்; மற்றபடி H_0-ஐ ஏற்றுக்கொள்கிறோம்.

இப்பொழுது $n_1+n_2=n$-ன் அளவு பெரியதாக இருந்தால், அதாவது $n > 20$ எனில் ஈருறுப்புப் பரவலானது தோராயமான இயல்நிலைப் பரவலாக அமைகிறது. α என்ற மிகைத்தன்மை மட்டத்தைப் பெற்ற தீர்வுகட்டமான வெளியானது, T-ன் எல்லா மதிப்புகளும் $T \leqslant t_1$ என்று இருப்பதற்கு ஏற்றதாக அமைகின்றது. இங்கு $t_1 = np^* + w_{\frac{\alpha}{2}} \sqrt{np^*(1-p^*)}$ ஆகும். இதே தீர்வுகட்டமான வெளியானது $T \geqslant t_2$ என்றிருப்பதற்குச் சமமாகும். இதில் $t_2 = np^* + w_{1-\frac{\alpha}{2}} \sqrt{np^*(1-p^*)}$ ஆகும்.

சோதனையில் T-ன் கண்டறிந்த மதிப்பானது t_2-ஐவிட அதிகமாக இருந்தாலோ அல்லது t_1-ஐவிடக் குறைவாகவோ அல்லது சமமாக இருந்தாலோ H_0-ஐ மறுக்கிறோம். மற்றபடி H_0-ஐ ஏற்றுக்கொள்ளுகிறோம்.

மாதிரி 1: சாரணர் பயிற்சி பெறும் 16 பள்ளி மாணவர்களுக்கு, அப் படையின் தலைவர், ஒரு சாட்டைக்கு ஒரே மாதிரியான முடியை X, Y என்ற இருவிதமான முறையில் போடுவதைப் பற்றிக் கற்றுக்கொடுக்கும் நோக்கத்தில், அம் மாணவர்களின் பாதியை ராண்டம் முறையில் தேர்ந்தெடுத்து, முதலில் முறை X-ஐயும் பின்பு முறை Y-ஐயும், மீதிப் பாதி மாணவர்களுக்கு முதலில் முறை Y-ஐயும், பின்பு முறை X-ஐயும் கற்பித்தார். சில மணி நேரங்கள் கழித்து அம் மாணவர்கள் ஒவ்வொருவருக்கும் சாட்டைகள் கொடுக்கப்பட்டு முடியைப் போடும்படி கேட்டுக்கொள்ளப்பட்டனர். அதன் விவரம் பின்வருமாறு குறிக்கப்பட்டது.

-	முடிபோடும் முறை		மொத்தம்
	X	Y	
அலைவெண்	14	2	16

சாட்டைக்கு முடியைப் போடுவதில் அம் மாணவர்கள் முறை X-ஐக் கையாளுவதற்கான நிகழ்தகவிற்கும், முறை Y-ஐக் கையாளுவதற்கான நிகழ்தகவிற்கு மிடையே வேறுபாடு உள்ளதா இல்லையா என்பதை ஈருறுப்புச் சோதனை மூலம் ஆய்க.

$\alpha = 0 \cdot 01$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். முறை X-ஐக் கையாளுவதற்கான நிகழ்தகவின் மதிப்பிற்கும், முறை Y-ஐக் கையாளுவதற்கான நிகழ்தகவின் மதிப்பிற்கும் இடையே வேறுபாடு இல்லை என்ற சூனிய எடுகோளைக் கொள்வோம். அதாவது $H_0: p_1 = p_2 = \frac{1}{2}$ ஆகும். முறை X-ஐக் கையாளுவதற்கான நிகழ்தகவு மதிப்பானது முறை Y-ஐக் கையாளுவதற்கான நிகழ்தகவின் மதிப்பைவிட அதிகமாக இருக்கும் என்ற மாற்று எடுகோளைக் கொள்வோம். இங்கு $n = 16$ ஆகும். மேலும், இங்குச் சிறிய அலைவெண் 2 ஆகும். எனவே, அட்டவணை D-லிருந்து $p(y \leqslant 2) = \cdot 002$ என்று அறிகிறோம். இந்த p-ன் மதிப்பான $\cdot 002$, $\alpha = \cdot 01$ என்ற மிகைத்தன்மை மட்டத்தைவிடக் குறைவாக உள்ளதால் H_0-ஐ $\alpha = \cdot 01$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

மாதிரி 2: இருவகைப்பட்ட தாவரங்களின் காய்கறி உற்பத்தியில், $\frac{1}{4}$ பங்கு மிகச் சிறியதாகவும், $\frac{3}{4}$ பங்கு பெரியதாகவும் இருப்பதாக எதிர்பார்க்கப்படுகிறது. இந்த எதிர்பார்ப்பை அறிவதற்கு நடத்திய ஒரு சோதனையில், 243 மிகச் சிறியனவாகவும், 682 பெரியனவாகவும் இருப்பதாகக் கண்டுபிடிக்கப்பட்டது. இப்பொழுது பெரிய காய்கறிகளை வகுப்பு 1 என்று குறிப்பிட்டால் $P^* = \frac{3}{4}$ ஆகும். மேலும் T என்பது பெரிய காய்கறிகளின் எண்ணிக்கையாகும்.

எனவே, சூனிய எடுகோள் $H_0 : p = \frac{3}{4}$ ஆகும்.

மாற்று எடுகோள் $H_1 : p \neq \frac{3}{4}$ ஆகும்.

இங்கு $n_1 + n_2 = n = 925$ ஆகும்.

$\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்தைக் கொண்ட தீர்வுகட்டமான வெளியை, அட்டவணை D-யிலிருந்து, பெரிய மாதிரிகளுக்குத் தோராயம் செய்வதன் மூலம் பெறலாம். எனவே, தீர்வுகட்டமான வெளியானது T-யின் எல்லா மதிப்புகளும் $T \leqslant t_1$ என்றிருப்பதற்கு ஏற்றதாக அமைகின்றது.

இங்கு $t_1 = np^* + w_{.025} \sqrt{np^* (1-p^*)}$

$$= (925)\tfrac{3}{4} + (-1{\cdot}960)\sqrt{(925)(\tfrac{3}{4})(\tfrac{1}{4})}$$
$$= \frac{2775}{4} + (-1{\cdot}960)\sqrt{\frac{2775}{16}}$$
$$= 693{\cdot}75 + (-1{\cdot}960)(\tfrac{1}{4})\sqrt{2775}$$
$$= 693{\cdot}75 - 0{\cdot}490\sqrt{2775}$$
$$= 693{\cdot}75 - 0{\cdot}490\,(52{\cdot}67)$$
$$= 693{\cdot}75 - 25{\cdot}80$$
$$= 667{\cdot}95$$

இதேபோன்று தீர்வுகட்டமான வெளியானது $T \geqslant t_2$ என்பதற்குச் சமமாகும்.

இதில் $t_2 = np^* + w_{.975} \sqrt{np^* (1-p^*)}$

$$= (925)\tfrac{3}{4} + (1{\cdot}960)\sqrt{(925)(\tfrac{3}{4})(\tfrac{1}{4})}$$
$$= 693{\cdot}75 + (1{\cdot}960)(\tfrac{1}{4})\sqrt{2775}$$
$$= 693{\cdot}75 + (0{\cdot}490)(52{\cdot}67)$$
$$= 693{\cdot}75 + 25{\cdot}80$$
$$= 719{\cdot}55$$

சோதனையில் T-ன் கண்டறிந்த மதிப்பு 682 ஆகும். இது t_1-ன் மதிப்பைவிட அதிகமாகவும், t_2-ன் மதிப்பைவிடக் குறைவாகவும் உள்ளதால் சூனிய எடுகோள் ஏற்றுக்கொள்ளப்படுகிறது.

தீர்வுகட்ட மட்டமான (critical level) $\hat{\alpha}$ ஐ இச் சோதனையில் அமைந்துள்ள ஏற்றுக்கொள்ளக்கூடிய வெளியானது (acceptance region), $np^* = 693{\cdot}75 = \dfrac{667{\cdot}95 + 719{\cdot}55}{2}$ என்ற மதிப்புக்கு இரு மருங்கிலும் அமைந்துள்ள ஏதோ ஒரு வெளி என்று கருதுவதால் பெறலாம். கண்டறிந்த மதிப்பான 682 ஆனது np^*-ஐவிடக் குறைவாக உள்ளதால் $\hat{\alpha}$-ன் பாதியை, சூனிய எடுகோள் உண்மையானது என்ற தற்கோளின் அடிப்படையில் $P(T \leqslant 682)$ என்று காண்பதன் மூலம் பெறலாம்.

$$\therefore\ P(T \leqslant 682) = P\left(\frac{T - np^*}{\sqrt{np^*(1-p^*)}} \leqslant \frac{682 - np^*}{\sqrt{np^*(1-p^*)}}\right)$$
$$\cong P\left(Z \leqslant \frac{682 - 693{\cdot}75}{52{\cdot}6}\right)$$

இங்கு Z என்பது திட்ட இயல்நிலைப் பரவலைப் (standard normal distribution) பெற்றுள்ளது.

எனவே, $\frac{\widehat{\alpha}}{2} = P(T \leqslant 682) \cong P(Z \leqslant -0{\cdot}223) \cong 0{\cdot}41.$

திட்ட இயல்நிலைப் பரவல் அட்டவணை T-யிலிருந்து பெறப் பட்டது. இவ்வாறாக, $\widehat{\alpha} = 2\,(0{\cdot}41) = 0{\cdot}82.$

குறைந்தது $0{\cdot}82$ என்ற மிகைத்தன்மை மட்டம் (level of significance) சூனிய எடுகோளான H_0-ஐ மறுப்பதற்குத் தேவைப் படுகிறது என்பதைக் காண்கிறோம். எனவே, கொடுக்கப்பட்ட விவரங்கள் சூனிய எடுகோளுக்கு மிகவும் பொருந்தியுள்ளன என்பதை அறிகிறோம்.

ஒருமுனைச் சோதனைக்கான எடுத்துக்காட்டுக் கணக்கு

ஓர் இயந்திரம் உதிரி உறுப்புகளை உற்பத்தி செய்கிறது. 5 சத வீதம் அல்லது அதற்குக் குறைவான சதவீதம் கொண்ட பழு தடைந்த அல்லது குறையுள்ள உறுப்புகளை உற்பத்தி செய்தால் இந்த இயந்திரம் நல்ல முறையில் வேலை செய்வதாகக் கருதப்படு கிறது. 5 சதவீத உறுப்புகளுக்கு மேல் குறையுள்ளதாக இருந் தால் இந்த இயந்திரம் பழுது செய்யப்பட வேண்டும் என்று கருதப்படுகிறது.

இப்பொழுது,

H_0 : இயந்திரம் நல்ல முறையில் வேலை செய்கிறது என்ற சூனிய எடுகோளை நாம் சோதனை செய்யும்போது,

H_1 : இயந்திரம் பழுதுபார்க்கப்பட வேண்டும் என்ற மாற்று எடுகோளைக் கருதுகிறோம்.

இயந்திரம் உற்பத்தி செய்யும் உதிரி உறுப்புகளில் ராண்டம் மாதிரியிலான 10 உதிரி உறுப்புகளைப் பரிசோதனை செய்வதன் மூலம் H_0-ஐச் சோதனை செய்வோம். ஒவ்வோர் உதிரி உறுப்பும் P என்ற ஒரே மாதிரியான நிகழ்தகவுடைய சார்பற்ற குறை யுள்ள உறுப்பு என்ற தற்கோளைக் (assumption) கொள்கிறோம்.

எனவே, H_0, H_1 என்ற மூல எடுகோள்கள்

$$H_0 : p \leqslant 0{\cdot}05$$
$$H_1 : p > 0{\cdot}05$$

என்ற எடுகோளுக்குச் சமமாகும்.

குறையுள்ள மொத்த உறுப்புகளின் எண்ணிக்கை T என்க. T ஆனது p, $n = 10$ என்ற சுட்டுறுப்புகளைக் கொண்ட ஈருறுப்புப் பரவலாக அமைந்துள்ளது. அட்டவணை D-லிருந்து H_0 உண்மையானால் $(p \leqslant 0{\cdot}05)$,

பிறகு, $P\ (T \leqslant 2) \leqslant 0{\cdot}9885$ ஆகும்.

$p = 0{\cdot}05$ எனில், $P\,(T < 2)$ ஆனது $0{\cdot}9885$-க்குச் சமமாக அமைகிறது.

மேலும் $P\,(T > 2) \leqslant 0{\cdot}0115$ ஆகும்.

$p = 0{\cdot}05$ எனில், இது $0{\cdot}0115$-க்குச் சமமாகும்.

T-ன் மதிப்பு 2-ஐவிட அதிகமாக இருந்தால் H_0-ஐ மறுப்பதாகத் தீர்மானிக்கிறோம்.

மாதிரி வெளியில் (sample space) அமைந்துள்ள புள்ளிக் கணமானது (set of points) $T > 2$ என்று இருக்குமாறு அமையும் T-ன் ஒவ்வொரு மதிப்புக்கும் ஏற்றவாறு அமைந்துள்ளது. இதனைத் தீர்வுகட்டமான வெளி (critical region) என்கிறோம். H_0 உண்மையாக இருக்கும்பொழுது தீர்வுகட்டமான வெளியில் இருக்குமாறு அமைவதற்கான ஒரு புள்ளியின் நிகழ்தகவு மிகவும் சிறியது $(< 0{\cdot}0115)$ என்பதால், தீர்மான விதியானது கீழ்க்கண்டவாறு அமைகிறது. கண்டறிந்த வெளிப்பாடு (observed outcome) தீர்வுகட்டமான வெளியில் $(T > 2)$ இருந்தால் H_0-ஐ மறுக்கிறோம். மற்றபடி H_0-ஐ ஏற்றுக்கொள்கிறோம்.

நம்பிக்கை இடைவெளி (Confidence Interval)

n சார்பற்ற முயற்சிகளைத் தழுவிய கண்டறிந்தவைகளைப் பெற்ற கூறு ஒன்றைப் பரிசோதனை செய்து Y என்ற ஒரு குறிப்பிட்ட நிகழ்ச்சியானது, நிகழும் எண்ணிக்கையைக் குறித்துக் கொள்வோம். n முயற்சிகளானவை ஒன்றை ஒன்று சார்பற்றன என்றும், ஒரு முயற்சிக்கும் அடுத்த முயற்சிக்குமிடையே ஏற்படும் நிகழ்ச்சியின் நிகழ்தகவான P நிலை எண்ணுகும் என்ற தற்கோளையும் கருதுகிறோம்.

$n \leqslant 30$ எனில், 95% அல்லது 99% நம்பிக்கை இடைவெளிகளுக்கான நம்பிக்கைக் கெழுக்களைத் தவிர, மற்ற நம்பிக்கைக் கெழுக்களைக் காண அட்டவணை D-ஐப் பயன்படுத்தலாம். தேவைப்படும் நம்பிக்கைக் கெழுவை $(1 - \alpha)$ என்று குறித்து $P_1 = \left(1 - \dfrac{\alpha}{2}\right)$ ஐக் கண்டு அட்டவணை D-யில் கூறு அளவான n-க்கு நிரையின்

குறுக்கே $y = (Y-1)$-க்கு ஏற்ற சரியான P_1 மதிப்பைக் கண்டு அதற்கேற்ற நிரலின் மேல் பகுதியில் P_1-ஐ உள்ளடக்கிய நிரலில் கண்ட p-ன் மதிப்பு L என்ற கீழ் எல்லையின் (lower limit) மதிப்பாகும்.

இப்பொழுது அடுத்த நிரையின் குறுக்கே $(y = Y)$-க்கு ஏற்ற $P_2 = \frac{\alpha}{2}$ ன் மதிப்புக்கு ஏற்ப நிரலின் மேல் பகுதியில் அமைந்துள்ள P_2-ஐ உள்ளடக்கிய நிரலில் கண்ட p-ன் மதிப்பு U என்ற மேல் எல்லையாக (upper limit) அமைகிறது. இவ்வாறு $n \leqslant 20$-க்கு நம்பிக்கை இடைவெளியைக் காண்கிறோம்.

$n > 20$-க்கு இயல்நிலைத் தோராயத்தைப் (normal approximation) பயன்படுத்தி,

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \sqrt{\frac{Y\ (n-Y)}{n^3}} \quad \text{and}$$

$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \sqrt{\frac{Y\ (n-Y)}{n^3}}$$

இங்கு $x_{1-\frac{\alpha}{2}}$ என்பது இயல்நிலைப் பரவலைப் பெற்ற ராண்டம் மாறியின் பகுதிமானம் (quantiles) ஆகும். இதனை இயல் நிலைப் பரவல் அட்டவணையான அட்டவணை T-யிலிருந்து பெறலாம்.

மாதிரி : கல்வியின் தரம் குறித்து, நாட்டுக் கல்விக்குழு அறிவித்துள்ள தகுதியை உயர்நிலைப் பள்ளிகள் பெற்றுள்ளனவா இல்லையா என்பதை அறியும் முகமாக 20 உயர்நிலைப் பள்ளிகள் ஒரு மாநிலத்திலிருந்து ராண்டம் முறையில் தேர்ந்தெடுக்கப் பட்டன. இதில் 7 உயர்நிலைப் பள்ளிகள் மட்டுமே 'சிறந்தது' என்ற தகுதியைப் பெற்றுள்ளனவாகக் காணப்பட்டன. மாநிலத்தி லுள்ள எல்லாப் பள்ளிகளையும் கருத்தில் கொள்ளும்போது இத் தகுதியைப் பெற்றுள்ள p என்ற விகிதப் பள்ளிகளுக்கான 95% நம்பிக்கை இடைவெளியைக் காண்க.

முதலில் மாநிலத்திலுள்ள உயர்நிலைப் பள்ளிகள் பெருமளவில் உள்ளதாகவும் அவைகளைச் சிறந்த தகுதியைப் பெற்றவை, சிறந்த தகுதியைப் பெறாதவை என்று பிரிக்கும்பொழுது ஒன்றையொன்று சார்பற்றனவாக இருக்குமாறும் கருதுகிறோம்.

2

இங்கு $n = 20$, $Y = 7$ ஆகும்.

இப்பொழுது $y = Y - 1 = 7 - 1 = 6$

அட்டவணை D-யிலிருந்து $n = 20$-க்கு $1 - \frac{\alpha}{2} = 1 - \frac{0 \cdot 05}{2}$ $= 0 \cdot 975$ நிகழ்தகவிற்கு ஏற்ற p-யின் மதிப்பானது, $p = 0 \cdot 15$-க்கு ஏற்ற $0 \cdot 9781$-க்கும், $p = 0 \cdot 20$-க்கு ஏற்ற $0 \cdot 9133$-க்குமிடையே அமைந்துள்ளது. இதுவே கீழ் எல்லையான (lower limit) L ஆகும்.

மேல் எல்லையைக் காண அட்டவணையில் அடுத்த நிரலில் $y = 7$-க்கும், $\frac{\alpha}{2} = \frac{0 \cdot 05}{2} = 0 \cdot 025$-க்கும் ஏற்ற p-யின் மதிப்பு $p = 0 \cdot 55$-க்கும், $p = 0 \cdot 59$-க்கும் இடையே அமைந்துள்ளது. இடைச் செருகல் செய்தால் $p = 0 \cdot 589 = U$ ஆகும்.

எனவே, $P(0 \cdot 15 < p < \cdot 589) = 0 \cdot 95$

கை - வர்க்க ஒரு கூறுச் சோதனை (Chi-square One Sample Test)

ஒரு குழுவிலுள்ளவர்களைப் பல்வேறு பிரிவுகளாக அமைத்து அப் பிரிவுகளுக்கிடையே அவைகளின் அலைவெண்கள் வேறுபடு கிறதா என்பதை இச் சோதனையின் மூலம் அறியலாம். உதாரண மாக ஒரு குழுவிலுள்ளவர்கள் ஏதோ ஒரு பொருள் குறித்து அவர் களின் அபிப்பிராயங்களைச் சாதகமாகவும், எதிர்த்தும், நடுநிலை வகித்தும் தெரிவிக்கலாம். எனவே, இக் குழு மூன்று பிரிவுகளாக அபிப்பிராய பேதத்தின் பேரில் பிரிக்கப்படுகிறது. இப்பொழுது இவ்வபிப்பிராயங்கள் அவைகளின் அலைவெண்களுக்கு வேறுபடு கிறதா என்பதை இச் சோதனையின் மூலம் அறியலாம். கைவர்க்கச் சோதனை இம் மாதிரியான பிரிவுபடுத்திய விவரங்களின் கண்டறிந்த அலைவெண்களுக்கும் எதிர்பார்க்கும் அலைவெண் களுக்கும் மிகைத்தன்மை செய்யுமளவுக்கு வேறுபாடு உள்ளதா என்பதை அறிய மிகவும் இன்றியமையாத சோதனையாக அமை கிறது. இதனை ஒரு பொருத்தச் செம்மைச் சோதனை (Goodness of fit test) என்றும் கூறலாம்.

கொடுக்கப்பட்ட கண்டறிந்த அலைவெண்கள், எதிர்பார்க்கும் அலைவெண்களுக்குப் போதுமான அளவு நெருங்கி உள்ளனவா என்பதை அறிய,

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

என்ற புள்ளியியல் அளவையைப் பயன்படுத்துகிறோம். இது $(k-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக

அமைகிறது. இதில் O_i என்பது i-ஆம் பிரிவிலுள்ள கண்டறிந்த அலைவெண்ணாகும். E_i என்பது H_0-ன் கீழ் i-ஆம் பிரிவிலுள்ள எதிர் பார்க்கும் அலைவு எண்ணாகும். கண்டறிந்த அலைவெண்களும், எதிர்பார்க்கும் அலைவெண்களும் ஓரளவு நெருங்கியிருப்பின் $(O_i - E_i)$-ன் மதிப்புச் சிறியதாக இருக்கும். இதனால் χ^2-ன் மதிப்பும் சிறியதாக இருக்கும். $(O_i - E_i)$-ன் மதிப்புப் பெரியதாக இருந்தால் χ^2-ன் மதிப்பும் பெரியதாக இருக்கும். எனவே, χ^2 மதிப்புப் பெரிய தாக இருப்பின் கண்டறிந்தவை ஒரே முழுமைத் தொகுதியி லிருந்து வரப்பெற்றனவாக இருக்க முடியாது என்று கிட்டத்தட்ட கொள்ளலாம்.

H_0-ன் கீழ் புள்ளியியல் அளவையான χ^2-ன் மாதிரிப் பரவ லானது $(k-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவ லாக அமைவதால் மேலுள்ள χ^2-ன் சூத்திரத்தின் மூலம் காணப் படும் கண்டறிந்த χ^2-ன் மதிப்பானது அட்டவணை C-ல் உள்ள $(k-1)$ வரையற்ற பாகைக்கும் α என்ற மிகைத்தன்மை மட்டத் திற்கும் ஏற்ற மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் H_0-ஐ α என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

மாதிரி: 42 விற்பனையாளர்கள் அவர்களின் விற்பனைத் திறன்களுக்கேற்ப மூன்று பிரிவுகளாகப் பிரிக்கப்பட்டு அவைகளுக் கான கண்டறிந்த அலைவெண்களும் எதிர்பார்க்கும் அலைவெண் களும் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டுள்ளன.

	நல்ல முறையிலான	திருப்தி கரமான	மோசமான	மொத்தம்
கண்டறிந்த அலைவெண்கள்	16	20	6	42
எதிர்பார்க்கும் அலைவெண்கள்	6·7	28·6	6·7	42

இப் பிரிவுகளின் எதிர்பார்க்கும் அலைவெண்களும், கண்டறிந்த அலைவெண்களும் மிகைத்தன்மை செய்யுமளவிற்கு வேறுபட்டு உள்ளனவா என்பதை கைவர்க்க ஒருகூறுச் சோதனையின் மூலம் ஆய்வோம்.

எதிர்பார்க்கும் அலைவெண்களுக்கும், கண்டறிந்த அலைவெண் களுக்கும் வேறுபாடு இல்லை என்ற சூனிய எடுகோளையும், எதிர்

பார்க்கும் அலைவெண்களுக்கும் கண்டறிந்த அலைவெண்களுக்கும் வேறுபாடு உள்ளது என்ற மாற்று எடுகோளையும் கொள்வோம். $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். இப்பொழுது,

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(16-6\cdot7)^2}{6\cdot7} + \frac{(20-28\cdot6)^2}{28\cdot6} + \frac{(6-6\cdot7)^2}{6\cdot7}$$

$$= \frac{(9\cdot3)^2}{6\cdot7} + \frac{(-8\cdot6)^2}{28\cdot6} + \frac{(-0\cdot7)^2}{6\cdot7}$$

$$= \frac{86\cdot49}{6\cdot7} + \frac{73\cdot96}{28\cdot6} + \frac{0.49}{6\cdot7}$$

$$= 12\cdot90 + 2\cdot59 + 0\cdot07$$

$= 15\cdot56$ ஆகும்.

இந்த χ^2 ஆனது $(k-1) = (3-1) = 2$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. எனவே, அட்டவணை C-யிலிருந்து $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்திற்கும் 2 வரையற்ற பாகைக்கும் ஏற்ற மதிப்பு $5\cdot99$ ஆகும். χ^2-ன் கண்டறிந்த மதிப்பான $15\cdot56$ ஆனது $5\cdot99$-ஐவிட அதிகமாக உள்ளதால் H_0-ஐ $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். அதாவது, கண்டறிந்த அலைவெண்களுக்கும், எதிர்பார்க்கும் அலைவெண்களுக்கும் வேறுபாடு உள்ளது என்று கொள்கிறோம்.

குறிப்பு: வரையற்ற பாகை = 1 எனில்; அதாவது $k = 2$ எனில் ஒவ்வோர் எதிர்பார்க்கும் அலைவெண்ணும் குறைந்தது 5 மதிப்பையாவது பெற்றிருக்க வேண்டும். வரையற்ற பாகை 1-க்கு மேலிருந்தால்; அதாவது $k > 2$ எனில், எதிர்பார்க்கும் அலைவெண்களில் 20 சதவீதம் 5-ஐ விடக் குறைவாக இருந்தாலோ அல்லது ஏதாவதோர் எதிர்பார்க்கும் அலைவெண் 1-ஐ விடக் குறைவாக இருந்தாலோ ஒரு கூறுக்கான கைவர்க்கச் சோதனையைப் பயன்படுத்தக்கூடாது: k-ன் அளவு பெரியதாக இருந்து எதிர்பார்க்கும் அலைவெண்களின் மதிப்பு 5 ஆகவோ அல்லது அதற்குக் குறைந்தோ இருந்தால், அவ்வலைவெண்களை அதற்கு முன்புள்ள அலைவெண்களோடு கூட்டியபின் ஒரு கூறுக்கான கைவர்க்கச் சோதனையைப் பயன்படுத்தலாம்.

'கோல்மோகோராவ்-ஸ்மிர்னுவின்' ஒரு கூறுச் சோதனை (Kolmogorov-Smirnov One Sample Test)

'கோல்மோகோராவ்-ஸ்மிர்னுவின்' ஒரு கூறுச் சோதனை ஒரு பொருத்தச் செம்மை (Goodness of fit) சோதனையாகும். அதாவது ஒரு கட்டான மாதிரி மதிப்புகளின் பரவலுக்கும், யாதானுமொரு குறிப்பிட்ட அறிமுறைக்குரிய (theoretical) பரவலுக்கும் உள்ள பொருத்தத்தின் படித்தரத்தைப் பற்றி (degree of agreement) அமைகிறது. இச் சோதனை மாதிரியின் மதிப்புகள் (scores) அறிமுறைக்குரிய பரவலைப் பெற்ற இனத் தொகுதியில் இருந்து வரப்பெற்றவைகளா என்பதைக் காண்கிறது. அதாவது, அறிமுறைக்கு உரிய பரவலின் கீழ் ஏற்படக்கூடிய குவிவ அலைவெண் பரவலைக் குறிப்பிட்டு அதனைக் கண்டறிந்த குவிவ அலைவெண் பரவலோடு ஒப்பிட்டுப் பார்ப்பதே இந்தச் சோதனையாகும்.

$X_1, X_2, \ldots, X_n$ என்ற n ராண்ட மாதிரிகளைப் பெற்ற விவரங்கள் ஏதோ ஒரு தெரியாத பரவல் சார்பான $F(x)$ஐப் பெற்று அமைவதாகக் கொள்வோம். மேலும் H_0-ன் கீழுள்ள எடுகோளுக்கான பரவல் சார்பான $F^*(x)$ஐத் தொடர்ச்சியானது என்று கொண்டால் சோதனையானது திருத்தமானது (exact) என்றும், இல்லையேல் திருத்தமற்றது என்றும் கொள்வோம். $F^*(x)$ ஆனது முற்றிலும் தெரிவிக்கப்பட்ட எடுகோள்மயமான பரவல் சார்பு (hypothesized distribution function) என்க. இப்பொழுது கீழ்க்காணும் எடுகோள்களைக் காண்போம் :

1. இருபுறச் சோதனை (Two sided test)

$-\infty$ லிருந்து $+\infty$ வரை உள்ள x-ன் எல்லா மதிப்புகளுக்கும்,

$$H_0 : F(x) = F^*(x)$$

குறைந்தபட்சம் x-ன் ஒரு மதிப்புக்காவது,

$$H_1 : F(x) \neq F^*(x)$$

2. ஒருபுறச் சோதனை (One sided test)

$-\infty$ லிருந்து $+\infty$ வரை உள்ள x-ன் எல்லா மதிப்புகளுக்கும்,

$$H_0 : F(x) \geqslant F^*(x)$$

குறைந்தபட்சம் x-ன் ஒரு மதிப்புக்காவது,

$$H_1 : F(x) < F^*(x)$$

3. ஒருபுறச் சோதனை (One sided test)

$-\infty$ லிருந்து $+\infty$ வரையுள்ள x-ன் எல்லா மதிப்புகளுக்கும்,

$H_0 : F(x) \leqslant F^*(x)$

குறைந்தபட்சம் x-ன் ஒரு மதிப்புக்காவது,

$H_1 : F(x) > F^*(x)$

$X_1, X_2, X_3, \ldots, X_n$ என்ற n ராண்டம் மாதிரிகளைத் தழுவிய $S(x)$ என்ற அனுபவப் பரவல் சார்பைக் (empirical distribution function) கருதுக. மேலுள்ள எடுகோள்களுக்கு ஏற்ற சோதனைப் புள்ளியியல் அளவை (Test Statistic) கீழ்க்கண்டவாறு அமைகிறது.

(1) T_1 என்ற புள்ளியியல் அளவையானது $S(x)$-க்கும் $F^*(x)$-க்கு மிடையேயுள்ள செங்குத்தான உச்ச தூரத்தைக் குறிப்பதாகக் கொள்வோம். அதாவது,

$$T_1 = \sup_x \left| F^*(x) - S(x) \right|$$

(2) T_1^+ என்ற புள்ளியியல் அளவையானது $S(x)$-க்கு மேல் $F^*(x)$ அடையும் உச்ச செங்குத்தான தூரம் என்க. அதாவது,

$$T_1^+ = \sup_x \left[F^*(x) - S(x) \right]$$

(3) T_1^- என்ற புள்ளியியல் அளவையானது $F^*(x)$-க்கு மேல் $S(x)$ அடையும் உச்ச செங்குத்தான தூரம் என்க. அதாவது,

$$T_1^- = \sup_x \left[S(x) - F^*(x) \right]$$

α என்ற மிகைத்தன்மை மட்டத்தில் (level of significance) T_1, T_1^+, T_1^- ஆகிய புள்ளியியல் அளவைச் சோதனைகள் $(1-\alpha)$க் கான பகுதி மானமான $w_{1-\alpha}$ன் மதிப்பை (அட்டவணை E) விட அதிகமாக இருந்தால் H_0-ஐ மறுக்கிறோம். அட்டவணை E ஆனது $n \leqslant 20$-க்கு, இருபுறச் சோதனையின் கீழ் துல்லியமானதாகும். ஒருபுறச் சோதனைக்கும் இருபுறச் சோதனையின் கீழ் $n > 20$-க்கும் இந்த அட்டவணை நல்ல தோராய மதிப்புகளைக் கொடுக்கிறது. $n > 40$-க்குத் தோராயமதிப்புகள் புள்ளியியல் அளவைகளுக்கான அணுகு கோடு பரவலைத் தழுவியதாக அட்டவணைப்படுத்தப் பட்டுள்ளது.

மாதிரி: X_1, X_2, X_3, ..., X_n என்ற ராண்டம் மாதிரிகளின் மதிப்புகள் பின்வருமாறு :

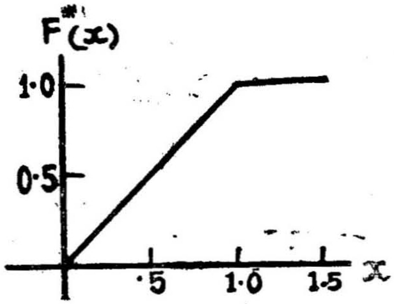
i	x_i
1	0·621
2	0·503
3	0·203
4	0·477
5	0·710
6	0·581
7	0·329
8	0·480
9	0·554
10	0·382

இம் மதிப்புகளால் வருவிக்கப்படும் பரவல் சார்பானது, சீரான பரவல் சார்பாக (uniform distribution) இருக்கும் என்ற சூனிய எடுகோளைக் கருதுக. அதாவது,

$$F^*(x) = 0 \quad , \quad x < 0 \text{ எனில்}$$

$$= x \quad , \quad 0 \leqslant x < 1 \text{ எனில்}$$

$$= 1 \quad , \quad 1 \leqslant x \text{ எனில்}$$



படம் 1

இப்பொழுது இதற்கேற்ற எடுகோள்கள் பின்வருமாறு:

சூனிய எடுகோள், $H_0 : F(x) = F^*(x), \forall x$

மாற்று எடுகோள், $H_1 : F(x) \neq F^*(x)$, (குறைந்தது ஒரு x-ன் மதிப்புக்காவது)

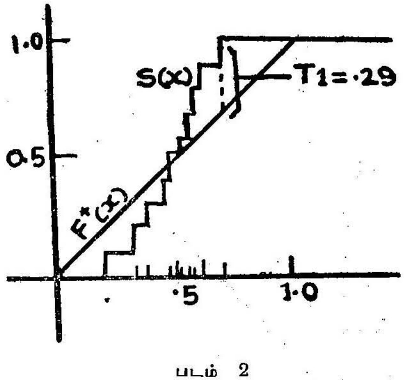
இதில் $F(x)$ என்பது எல்லா X_i-களுக்கும் பொதுவான தெரியாத பரவல் சார்பாகும்.

$F^*(x)$ என்பது படம் (1)-ல் உள்ளதுபோல் ஆகும். இங்கு கோல்மோகோராவின் ஒருபுறச் சோதனையைப் பயன்படுத்துகிறோம்.

$\alpha = 0{\cdot}05$ என்ற தீர்வுகட்டமான வெளியின் அளவு (critical region of size, α) T_1-ன் மதிப்புகளுக்கு, அதாவது $0{\cdot}95$ பகுதிமான மதிப்பான $0{\cdot}409$-க்கு ($n = 10$-க்கு அட்டவணை E-லிருந்து) அதிகமான மதிப்புக்குப் பொருந்தியதாகும்.

T_1-ன் மதிப்பை $F^*(x)$-க்கு மேல் $S(x)$ என்ற அனுபவப் பரவல் சார்பை வரைபடம் மூலம் அமைத்து இவைகளுக்கிடையேயுள்ள செங்குத்தான தூரத்தை அளப்பதன் மூலம் பெறலாம். இதனைப் படம் (2)-ல் காண்போம்.

இவ்வாருக $x = 0{\cdot}710$ என்னுமிடத்தில் T_1-ன் மதிப்பு $0{\cdot}290$ ஆகும்.



படம் 2

$$S\ (0{\cdot}710) = 1{\cdot}000$$
$$F^*(0{\cdot}710) = 0{\cdot}710$$

எனவே, $T_1 = \sup\limits_{x} \left| F^*(x) - S(x) \right|$

$= \left| F^*(\cdot 710) - S(\cdot 710) \right|$

$= 1\cdot 000 - 0\cdot 710 = 0\cdot 290$

$T_1 = 0\cdot 290 < 0\cdot 409$ என்பதால் சூனிய எடுகோளை ஏற்றுக்கொள்ளுகிறோம்.

அட்டவணை E-யிலிருந்து தீர்வுகட்டமான மட்டம் $\widehat{\alpha}$ ஆனது, $0\cdot 20$-ஐ விட அதிகமாக இருப்பதைக் காண்கிறோம். இப்பொழுது,

$H_0 :\ F(x) \geqslant F^*(x)\ \forall\ x$ என்று கொண்டால்

$H_1 :\ F(x) < F^*(x)$, ஏதோ x-க்கு

என்ற ஒருபுற மாற்று எடுகோளை அமைத்துக் கொள்கிறோம். $\alpha = 0\cdot 05$ என்ற மிகைத்தன்மை மட்டத்தில் H_0-ஐ ஒருபுறச் சோதனையின் கீழ் T_1-ன் மதிப்பு $0\cdot 95$ பகுதிமான மதிப்புக்கு அதிகமாக இருந்தால் மறுக்கிறோம். அவ்வாறே அட்டவணை E-யிலிருந்து, $n = 10$-க்கு இங்கு $0\cdot 95$-ன் பகுதிமான மதிப்பு $0\cdot 369$ ஆகும்.

T_1^+-ன் மதிப்பு $= \sup\limits_{x} \left[F^*(x) - S(x) \right]$

$= F^*(0\cdot 3289) - S(0\cdot 3289)$

$= 0\cdot 3289 - 0\cdot 100$

$= 0\cdot 2289$ ஆகும்.

ஒருபுறச் சோதனையின் கீழ் மாற்று திசைக்கு

$T_1^- = \sup\limits_{x} \left[S(x) - F^*(x) \right]$

$= S(0\cdot 710) - F^*(0\cdot 710)$

$= 1\cdot 000 - 0\cdot 710$

$= 0\cdot 290$

இங்கு இருபுறச் சோதனைதான் மிகவும் ஏற்புடையதான சோதனையாக அமைந்துள்ளது. ஒருபுறச் சோதனையை எவ்வாறு பயன்படுத்தி T_1^+, T_1^- ஐக் காண்பது என்பதற்காகவே ஒருபுறச் சோதனை பயன்படுத்தப்பட்டது.

மாதிரி: கருநிறத்திலிருந்து சுமாரான மாநிறம் வரையுள்ள படங்கள் தரப்படுத்தி அமைக்கப்பட்டு (ranked), கண்டறிந்த பரவலின் மதிப்புகளுக்கும் அறிமுறைக்குரிய பரவல் மதிப்புகளுக்கும் இடையே வேறுபாடுகள் இல்லை என்ற சூனிய எடுகோளின் கீழ் அறிமுறைக்கான குவிவப் பரவலும், கண்டறிந்த விவரங்களுக்கு (observed choices) ஏற்ற குவிவப் பரவலும் கீழ்க் கண்டவாறு அட்டவணைப்படுத்தப்பட்டன.

தரம் (ranks)	1	2	3	4	5
f: தரத்தைப் பெற்ற நபர்களின் எண்ணிக்கை	0	1	0	5	4
$F^*(x)$: H_0-ன் கீழ் அமைந்துள்ள அறிமுறைக்குரிய பரவல்	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{5}{5}$
$S(x)$: கண்டறிந்த விவரங்களுக்கேற்ற பரவல்	$\frac{0}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{6}{10}$	$\frac{10}{10}$
$T_1 = \mid F^*(x) - S(x) \mid$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{5}$	0

மேலுள்ள அட்டவணையில்,

$T_1 = \mid F^*(x) - S(x) \mid$-ன் உச்ச மதிப்பு $\frac{1}{2}$ ஆகும். அட்டவணை E-லிருந்து, $n = 10$-க்கு 0·95 பகுதிமானத்தின் மதிப்பு 0·489 ஆகும்.

$T_1 = 0{\cdot}5 > 0{\cdot}489$ என்பதால் H_0-ஐ மறுக்கிறோம்.

ஒரு கூறு ஓட்டச் சோதனை (One Sample Run Test)

ஒரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட ராண்டம் மாதிரிகள் ராண்டம் முறையில் அமைந்துள்ளனவா இல்லையா என்பதை அறிய இச் சோதனை பயன்படுகிறது. தொடர்ச்சியாக ஒரு குறிப்பிட்ட நிகழ்ச்சி வேறுபட்ட பிறிதொரு நிகழ்ச்சி நிகழும்வரை நிகழ்ந்தால் அது ஒரு ஓட்டம் என்று கொள்ளப்படுகிறது.

உதாரணமாக ஒரு நாணயம் 10 முறை சுண்டப்படுகிறது. அதன் விளைவாக *H H H T T H H T T T* என்ற நிகழ்ச்சிகள் பதிவு

செய்யப்பட்டன என்க. இதில் H ஆனது தொடர்ச்சியாக மூன்று முறை நிகழ்ந்துள்ளது. பின்பு T ஆனது 2 முறை நிகழ்ந்துள்ளது. மறுபடியும் H ஆனது இருமுறையும் T மூன்று முறையும் நிகழ்ந்துள்ளன.

எனவே, இங்கு $\underline{HHH}$ $\underline{TT}$ $\underline{HH}$ $\underline{TTT}$ என்ற நான்கு ஓட்டங்கள் நிகழ்ந்துள்ளன. ஓட்டங்களின் எண்ணிக்கையை r என்று குறித்தால், இங்கு $r = 4$ ஆகும். n_1 என்பது H-களின் எண்ணிக்கைகள் என்றும், n_2 என்பது T-களின் எண்ணிக்கைகள் என்றும் கொண்டால் $N = n_1 + n_2$ ஆகும். ஒரு கூறுச் சோதனையைப் பயன்படுத்த முதலில் n_1, n_2 ஆகியவைகளின் நிகழ்ச்சிகள் நிகழும் அதே வரிசையில் ஏற்படும் ஓட்டங்களின் எண்ணிக்கையான r-ஐக் காணவேண்டும். n_1, n_2 ஆகியவைகளின் எண்ணிக்கை 20-ஐ விடக் குறைவாகவோ அல்லது சமமாகவோ இருந்தால் அட்டவணை F-ஐப் பயன்படுத்தி H_0-ன் கீழ் $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் r-க்கான தீர்வு மட்டமான மதிப்புகளைப் (critical values) பெறமுடியும். இவை r-ன் மாதிரிப் பரவலிலிருந்து (sampling distribution) பெறப்பட்ட தீர்வு கட்டமான மதிப்புகளாகும். r-ன் கண்டறிந்த மதிப்புகள் தீர்வு கட்டமான மதிப்புகளுக்கிடையே அமைந்தால் H_0-ஐ ஏற்றுக் கொள்ளுகிறோம். r-ன் கண்டறிந்த மதிப்புகள் தீர்வு கட்டமான மதிப்பின் எல்லை மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் H_0-ஐ மறுக்கிறோம். அட்டவணை F_1-ல் r-ன் மிகச் சிறிய மதிப்புகளுக்கு அவை H_0-ன் கீழ் ஏற்படக்கூடிய நிகழ்தகவு $p = 0{\cdot}025$ என்று இருக்குமாறு கொடுக்கப்பட்டுள்ளது.

அட்டவணை F_2-ல் r-ன் மிகப் பெரிய மதிப்புகளுக்கு அவை H_0-ன் கீழ் ஏற்படக்கூடிய நிகழ்தகவு $p = 0{\cdot}025$ என்றிருக்குமாறு அட்டவணைப்படுத்தப்பட்டுள்ளது.

கண்டறிந்த r-ன் மதிப்பு அட்டவணை F_1-லுள்ள மதிப்பிற்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால் அல்லது அட்டவணை F_1-லுள்ள மதிப்பிற்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கத்தகும் வெளியில் (rejection region) இருக்கும்.

உதாரணமாக, ஒரு நாணயத்தை 10 முறை சுண்டியதில் $r = 4$, $n_1 = 5$, $n_2 = 5$; அதாவது ஓட்டங்கள் 4-ம், H-கள் 5-ம், T-கள் 5-ம் என்று பெறப்பட்டன. அட்டவணை F-ல் $n_1 = 5$, $n_2 = 5$-க்கு ஒரு ராண்டம் மாதிரி, 2 ஓட்டங்களுக்கு மேல்; ஆனால் 10 ஓட்டங்களுக்குக் குறைவாகப் பெற்றிருக்க வேண்டும்

என்று காட்டுகிறது. கண்டறிந்த r-ன் மதிப்பு 2-ஐ விடக் குறைவாகவோ அல்லது 10-க்குச் சமமாகவோ அல்லது 16-க்கு அதிகமாகவோ இருந்தால், அது, $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கத்தகும் வெளியில் இருக்கும். இங்குக் கண்டறிந்த r-ன் மதிப்பு = 4 ஆகும். இது 2-ஐ விட அதிகமாக உள்ளது. எனவே $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் நாணயமானது ராண்டம் தொடர்ச்சியை ஏற்படுத்துகிறது என்ற சூனிய எடுகோளை ஏற்றுக்கொள்கிறோம்.

ஒரு முனைச் சோதனையைப் பயன்படுத்தவேண்டுமென்றால், கொடுக்கப்பட்டுள்ள இரு அட்டவணைகளில் ஏதாவது ஒன்றைப் பயன்படுத்தினால் போதும். உதாரணமாக $r = 15, n_1 = 5, n_2 = 5$ என்க. இங்குக் கண்டறிந்த r-ன் மதிப்பு 15 ஆகும். இது F_2 அட்டவணையில் $n_1 = 5, n_2 = 6$-க்கு ஏற்ற மதிப்பான 10-ஐ விட அதிகமாக இருப்பதால் H_0-ஐ $\alpha = 0{\cdot}025$ என்ற மிகைத் தன்மை மட்டத்தில் மறுக்கிறோம்.

மாதிரி: ஒரு பரிசோதனையில் 24 சிறுவர்களின் ஆக்கிரமிப்புத் தன்மை குறித்து ஆய்வு செய்யும் நோக்கத்தில் ஜோடியாகச் சிறுவர்கள் விளையாடும் நிகழ்ச்சிகள் கண்காணிக்கப்பட்டன. இதில் பெரும்பாலான சிறுவர்கள் ஒரே பள்ளியிலிருந்து வந்துள்ளனர் என்று அறிந்தபின், இந்தச் சோதனையின் ஆய்வு ஒரு சார்பானதாகவோ (biased) அல்லது ராண்டம் முறையில் அமையாமலோ இருக்கும் என்று கருதப்பட்டது. இதனை அறிய அச் சிறுவர்களின் ஆக்கிரமிப்புக்கான மதிப்புகள் அக் குழுவின் இடைநிலையின் மதிப்பைவிட அதிகமாக இருந்தால் + குறியுடைமை, குறைவாக இருந்தால் − குறியுடைமை எனக் குறிக்கப்பட்டு அட்டவணைப்படுத்தப்பட்டு (அட்டவணை பக்கம் 29 கண்க), ஒரு கூறு ஓட்டச் சோதனை பயன்படுத்தப்பட்டது.

சூனிய எடுகோளான H_0-ஐ +, − ஆகிய குறிகள் ராண்டம் முறையில் நிகழ்ந்துள்ளன என்றும், மாற்று எடுகோளான H_1-ஐ, அக் குறிகள் ராண்டம் முறையைத் தழுவவில்லை என்றும் கொள்வோம்.

இங்கு $N = 24$. இடைநிலைக்கேற்ற குறிகளைக் கருத்தில் கொண்டால்,

$$\underbrace{+}_{1}\ \underbrace{-}_{2}\ \underbrace{++++}_{3}\ \underbrace{-}_{4}\ \underbrace{+++}_{5}\ \underbrace{-----}_{6}\ \underbrace{+}_{7}\ \underbrace{--}_{8}\ \underbrace{+++}_{9}\ \underbrace{----}_{10}$$

சிறுவர்கள்	மதிப்புகள்	இடைநிலைக் கேற்ற +, − குறிகள்
1	42	+
2	34	−
3	47	+
4	54	+
5	62	+
6	55	+
7	23	−
8	37	+
9	54	+
10	86	+
11	13	−
12	14	−
13	26	−
14	29	−
15	89	+
16	35	−
17	24	−
18	38	+
19	97	+
20	72	+
21	24	−
22	18	−
23	17	−
24	19	−

$r = 10$

$n_1 = +$ குறிகளின் எண்ணிக்கை 12

$n_2 = -$ குறிகளின் எண்ணிக்கை 12

அட்டவணை F_1-ல் $n_1 = 12, n_2 = 12$-க்கான மதிப்பு 7 ஆகும். F_2-ல் 19 ஆகும். கண்டறிந்த r-ன் மதிப்பான 10, F_1-அட்டவணையின் மதிப்பான 7-ஐ விட அதிகமாக உள்ளது. அல்லது, F_2 அட்டவணையின் மதிப்பான 19-ஐ விடக் குறைவாக உள்ளது. எனவே, கண்டறிந்த r-ன் மதிப்பு ஏற்றுக்கொள்ளும் வெளியில் அமைவதால் H_0-ஐ ஏற்றுக்கொள்ளுகிறோம். அதாவது மாதிரியின் மதிப்புகள் ராண்டம் முறையில் அமைந்துள்ளன என்பதாகும்.

கவனிக்கவும் : n_1-ன் அளவோ அல்லது n_2-ன் அளவோ 20-ஐ விடப் பெரியதாக இருந்தால், கருத்தில் கொள்ளப்பட்ட மாதிரி பெருமளவிலான மாதிரி (large sample) என்று கருதப்படுகிறது.

அம்மாதிரியான நிலையில் r-ன் மாதிரிப் பரவலுக்கு, ஒரு நல்ல தோராயமான இயல்நிலைப் பரவலைக் கருதுவதால்,

$$\text{சராசரி } \mu_r = \left\{ \frac{2n_1\, n_2}{n_1+n_2} \right\} + 1$$

$$\text{திட்ட விலக்கம்} = \sigma_r = \sqrt{\frac{2n_1\, n_2\, (2n_1\, n_2 - n_1 - n_2)}{(n_1+n_2)^2\, (n_1+n_2-1)}}$$

எனவே, n_1 அல்லது n_2-வோ 20-ஐ விட அதிகமாக இருந்தால், H_0-ஐ

$$Z = \frac{r - \mu_r}{\sigma_r}$$

$$= \frac{r - \left\{ \dfrac{2n_1\, n_2}{n_1+n_2} - 1 \right\}}{\sqrt{\dfrac{2n_1\, n_2\, (2n_1\, n_2 - n_1 - n_2)}{(n_1+n_2)^2\, (n_1+n_2-1)}}}$$

என்ற புள்ளியியல் அளவை மூலம் பரிசோதனை செய்கிறோம். Z ஆனது தோராயமாக $N\,(0,1)$ ஆக இருப்பதால் கண்டறிந்த Z-ஐ இயல்நிலைப் பரவல் அட்டவணை A-ஐப் பார்ப்பதன் மூலம் மிகைத்தன்மையை (significance) அறியலாம். அட்டவணை A-ல் ஒருமுனை நிகழ்தகவு மதிப்புகள் கொடுக்கப்பட்டுள்ளன.

மாதிரி: ஓர் அஞ்சல் அலுவலகத்தில் அஞ்சல் தலைகள் கொடுக்கும் இடத்தில் நின்றுகொண்டிருக்கும் 50 நபர்கள் ஆண்களும் பெண்களுமாகப் பின்வருமாறு நின்றுகொண்டிருந்தனர். இந்த வரிசை ராண்டம் மாதிரியில் அமைந்துள்ளதா இல்லையா என்பதை ஒருகூறு ஓட்டச் சோதனை மூலம் காண்க.

M F M F MMM FF M F M F
M F MMMM F M F M F MM
FFF M F M F M F MM F
MM F MMMM F M F MM

இப்பொழுது மேலுள்ள விவரத்திற்கு உரிய ஓட்டங்களைக் காண்போம்.

M F M F MMM FF M F M F
M F MMMM F M F M F MM
FFF M F M F M F MM F
MM F MMMM F M F MM

இதில் $r = 35$

n_1 = ஆண்களின் எண்ணிக்கை 30

n_2 = பெண்களின் எண்ணிக்கை 20

$r \geqslant 35$ ஓட்டங்கள் H_0-ன் கீழ் நிகழ்வதற்கான வாய்ப்பு உள்ளதா என்பதை அறிய,

$Z = \dfrac{r - \mu_r}{\sigma_r}$ என்ற புள்ளியியல் அளவையைக் கருதுகிறோம்.

$$\text{இதில் } \mu_r = \left\{\frac{2\, n_1\, n_2}{n_1 + n_2}\right\} + 1$$

$$= \left\{\frac{2\,(30)\,(20)}{30 + 20}\right\} + 1$$

$$= \left\{\frac{1200}{50}\right\} + 1$$

$$= \frac{1200 + 50}{50}$$

$$= \frac{1250}{50}$$

$$= 25$$

$$\sigma_r = \sqrt{\frac{2\, n_1 n_2\, (2\, n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2\, (n_1 + n_2 - 1)}}$$

$$= \sqrt{\frac{2\,(30)\,(20)\,(2 \times 30 \times 20 - 30 - 20)}{(30 + 20)^2\,(30 + 20 - 1)}}$$

$$= \sqrt{\frac{1200\,(1200 - 50)}{(50)^2\,(49)}}$$

$$= \sqrt{\frac{1200 \times 1150}{2500 \times 49}}$$

எனவே, $Z = \frac{r - \mu_r}{\sigma_r}$

$$= \frac{35 - 25}{3 \cdot 3574}$$

$$= \frac{10}{3 \cdot 3574} = 2 \cdot 98$$

அட்டவணை A-யிலிருந்து H_0-ன் கீழ் $Z \geqslant 2 \cdot 98$ என்று ஏற்படுவதற்கான நிகழ்தகவு $p = 2\,(0 \cdot 0014) = 0 \cdot 0028$ (ஏனெனில் அட்டவணை A-ல் ஒருமுனை நிகழ்தகவு மதிப்புகள் மட்டுமே கொடுக்கப்பட்டுள்ளன).

எனவே, $p = 0 \cdot 0028$ ஆனது $\alpha = 0 \cdot 05$ என்ற மிகைத் தன்மை மட்டத்தைவிடக் குறைவாக உள்ளதால் H_0-ஐ மறுக்கிறோம். அதாவது 50 நபர்கள் நின்றுகொண்டிருந்த வரிசையானது ராண்டம் முறையில் அமையவில்லை என்பதேயாகும்.

'லில்லிஃபார்ஸ்' சோதனை (Lilliefors Test)

$X_1, X_2, X_3, \ldots, X_n$ என்ற ராண்டம் மாதிரியானது $F(x)$ என்ற தெரியாத பரவல் சார்பைத் தழுவி இருப்பதாகக் கொள்வோம். மாதிரிச் சராசரியான $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$-ஐக் கணக்கிடுவோம். இதனை μ-க்கு மதிப்பீடாகக் கருதுகிறோம்.

மேலும் $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}$ ஐ, σ-க்கு மதிப்பீடாகக் கருதிக் கணக்கிடுவோம்.

இப்பொழுது $Z_i = \frac{X_i - \bar{X}}{s}$, $i = 1, 2, 3, \ldots\ldots, n$ என்ற இயல்நிலையாக அமைந்த மாதிரி மதிப்புகளைக் காண்கிறோம்.

இச் சோதனையைச் செய்வதற்கு முன்பு, மாதிரியானது ராண்டம் மாதிரி என்ற தற்கோளையும் (assumption), தெரிவிக்கப்படாத சராசரி, மாறுபாடு ஆகியவைகளைப் பெற்ற இயல்நிலைப் பரவலாக ராண்டம் கூறுகள் அமைந்துள்ளன என்ற சூனிய எடுகோளையும், X_i-களின் பரவல் சார்பு இயல்நிலைப் பரவலாக அமையவில்லை என்ற மாற்று எடுகோளையும் கொள்கிறோம்.

சாதாரணமாக இங்குப் புள்ளியியல் சோதனையானது வழக்க மான இருபுற கோல்மோகோராவ் சோதனைப் புள்ளியியல் அளவை யாகும். இப் புள்ளியியல் அளவையானது X_i-களின் அனுபவப் பரவல் சார்பிற்கும், சராசரி $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$, திட்ட விலக்கம் $s = \sqrt{\frac{1}{n-1} \sum_{1}^{n} (X_i - \bar{X})^2}$ ஐப் பெற்ற இயல்நிலைப் பரவலிற்கு மிடையே உள்ள உச்ச செங்குத்தான தூரமாகும். இப் புள்ளியியல் அளவையை எளிதாகக் கீழ்க்கண்டவாறும் காணமுடியும்.

அட்டவணை T-ஐப் பயன்படுத்தி திட்ட இயல்நிலைப் பரவல் சார்புக்கு வரைபடம் ஒன்றை வரைந்து அதனை $F^*(x)$ எனக் குறிப்பிடவும். மேலும், இயல்நிலையாக்கப்பட்ட $S(x)$ என்ற மாதிரியின் (normalised sample) அனுபவப் பரவல் சார்பை (empirical distribution function) வரைபடம் மூலம் அமைக்கவும். அதாவது Z_i-களுக்கு அனுபவப் பரவல் சார்பை அமைக்க வேண்டும். $F^*(x)$-க்கும் $S(x)$-க்கும் இடையேயுள்ள உச்ச குத்துயரத்தைச் சோதனைப் புள்ளியியல் அளவையாகக் கருது கிறோம். அதாவது, லில்லிஃபார்ஸ் சோதனைக்கான புள்ளியியல் அளவை,

$$T_2 = \sup_{x} | F^*(x) - S(x) |$$ ஆகும்.

இச் சோதனையில் T_2 என்ற புள்ளியியல் அளவையைக் காணும் போது $S(x)$ இயல்நிலைப் படுத்தப்பட்ட மாதிரியில் இருந்து பெறப்பட்டதாகும். ஆனால், கோல்மோகோராவ் சோதனையில் T_1-ஐக் காணும்போது $S(x)$-ஐக் கொடுக்கப்பட்ட கண்டறிந்தவை களிலிருந்து பெறுகிறோம்.

T_2-ன் மதிப்பானது α என்ற மிகைத்தன்மை மட்டத்தைக் கருதியபின் அட்டவணை U-லிருந்து கண்ட $(1-\alpha)$-க்கான பகுதி மான மதிப்பைவிட அதிகமாக இருந்தால் H_0-ஐ மறுக்கிறோம்.

மாதிரி : தொலைபேசி எண்களைக் கொண்ட புத்தகத்தில் இருந்து 50 இரட்டைப்படை எண்கள் ராண்டம் முறையில் தேர்ந் தெடுக்கப்பட்டன. இந்த எண்கள் இயல்நிலைப் பரவலைப் பெற்ற ராண்டம் மாறிகளில் இருந்து எடுக்கப்பட்ட கண்டறிந்தவை

களா என்பதை அறியும்பொருட்டு லில்லிஃபார்ஸ் சோதனை செய்யப்பட்டது. இந்த எண்கள் வரிசைப்படுத்தப்பட்டுக் கீழ்க் கண்டவாறு அமைக்கப்பட்டன.

23, 23, 24, 27, 29, 31, 32, 33, 33, 35,
36, 37, 40, 42, 43, 43, 44, 45, 48, 48,
54, 54, 56, 57, 57, 58, 58, 58, 58, 59,
61, 61, 62, 63, 64, 65, 66, 68, 68, 70,
73, 73, 74, 75, 77, 81, 87, 89, 93, 97.

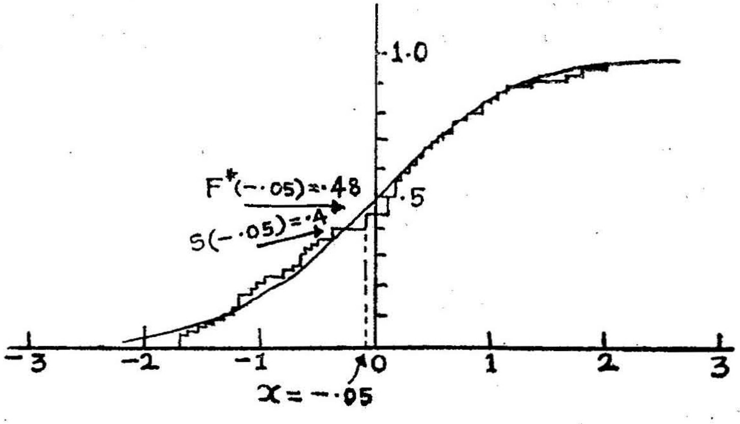
இந்த எண்கள் இயல்நிலைப் பரவலைப் பெற்ற ராண்டம் மாறி களின் கண்டறிந்தவைகளாகும் என்ற சூனிய எடுகோளையும், இயல்நிலைப் பரவலைப் பெற்ற மாறிகளின் கண்டறிந்தவைகளல்ல என்ற மாற்று எடுகோளையும் கொள்வோம். இப்பொழுது கொடுக்கப்பட்ட கண்டறிந்தவைகளின் மதிப்பை $Z = \frac{X_i - \bar{X}}{s}$ என்ற சூத்திரத்தைப் பயன்படுத்தி இயல்நிலைப்படுத்திக் கீழ்க்கண்ட அட்டவணையில் அமைப்போம்.

X_i	Z_i	X_i	Z_i
23	−1·69	43	−0·63
23	−1·69	44	−0·58
24	−1·63	45	−0·53
27	−1·48	48	−0·37
29	−1·37	48	−0·37
31	−1·27	54	−0·05
32	−1·21	54	−0·05
33	−1·16	56	0·05
33	−1·16	57	0·10
35	−1·05	57	0·10
36	−1·00	58	0·16
37	−0·95	58	0·16
40	−0·79	58	0·16
42	−0·69	58	0·16
43	−0·63	59	0·21

X_i	Z_i	X_i	Z_i
61	0·31	73	0·95
61	0·31	73	0·95
62	0·37	74	1·00
63	0·42	75	1·05
64	0·47	77	1·16
65	0·52	81	1·37
66	0·58	87	1·68
68	0·68	89	1·79
68	0·68	93	2·00
70	0·79	97	2·21

இப்பொழுது 'லில்லிஃபார்ஸ்' சோதனைக்கான புள்ளியியல் அளவை, $T_2 = \sup_x | F^*(x) - S(x) |$ ஆகும். இதில் $F^*(x)$ என்பது திட்ட இயல்நிலைப் பரவல் சார்பாகும். $S(x)$ என்பது Z_i-களின் அனுபவப் பரவலாகும். $F^*(x)$-க்கும், $S(x)$-க்கும் உச்ச குத்துயரத் தூரம் கீழ்க்காணும் படத்தில் அமைக்கப்பட்டுள்ளது. படத்தில் $F^*(x)$-க்கும், $S(x)$-க்கும் உள்ள உச்ச குத்துயரத் தூரம்,

$$F^*(x) - S(x) = 0{\cdot}48 - 0{\cdot}40 = 0{\cdot}08 \text{ ஆகும்.}$$



$\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்தில் $w_{0 \cdot 95}$ என்ற பகுதிமான மதிப்பை அட்டவணை U-ல் இருந்து கண்டால், அதன் மதிப்பு

$$w_{0 \cdot 95} = \frac{0 \cdot 886}{\sqrt{n}}$$

$$= \frac{0 \cdot 886}{\sqrt{50}}$$

$$= \frac{0 \cdot 886}{7 \cdot 071}$$

$$= 0 \cdot 125 \text{ ஆகும்.}$$

$T_2 = 0 \cdot 08 < w_{0 \cdot 95} = 0 \cdot 125$ என்பதால், H_0-ஐ $\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்தில் ஏற்றுக்கொள்ளுகிறோம்.

'க்ராமர்-வான் மைஸலின்' பொருத்தச் செம்மைக்கான சோதனை (The Cramer - von Mises Goodness Fit Test)

தெரிவிக்கப்படாத $F(x)$ என்ற பரவல் சார்புடன் தொடர்புடைய $X_1, X_2, X_3, \ldots\ldots, X_n$ என்ற மாதிரிக் கூறு ஒன்றைக் கருதுக. இதனை வரிசைப்படுத்திக் கிடைத்த ராண்டம் மாதிரியை $X^{(1)} \leqslant X^{(2)} \leqslant \ldots\ldots\ldots \leqslant X^{(n)}$ என்று குறிப்போம். இதில் $X^{(i)}$ என்பது தரம் (rank) i-ஐப் பெற்ற வரிசைப் புள்ளியியல் அளவையாகும் (order statistics). இப்பொழுது $F^*(x)$ என்பது முற்றிலும் தெரிவிக்கப்பட்ட எடுகோள் மயமான பரவல் சார்பு (hypothesised distribution function) என்க. இச் சோதனையில் $F^*(x)$ தொடர்ச்சியானது என்று கருதுகிறோம். $-\infty$ லிருந்து $+\infty$ வரை எல்லா x-க்கும் $F(x) = F^*(x)$ என்ற சூனிய எடுகோளையும், குறைந்தது ஒரு x-க்காவது $F(x) \neq F^*(x)$ என்ற மாற்று எடுகோளையும் கொள்கிறோம்.

ராண்டம் மாதிரியில் கண்டறியும் i-ஆம் சிறிய எண்ணுக்கு ஏற்ற எடுகோள் மயமான பரவல் சார்பு $F^*(x)$-ன் மதிப்பு $F^*\left(X^{(i)}\right)$ என்க. $i = 1, 2, \ldots\ldots n$.

இச் சோதனையில் புள்ளியியல் அளவையை

$$T_3 = \frac{1}{12n} + \sum_{i=1}^{n} \left[F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right)\right]^2$$

என்று குறிப்பிடுகிறோம்.

T_3-ன் பகுதி மானங்களை (quantiles) T_3-ன் அணுகு கோடு பரவல் சார்பைப் பயன்படுத்தித் தோராயமாக ஆண்டர்ஸன், டார்லிங் ஆகியோர் அட்டவணைப்படுத்தியுள்ளனர். அவைகளாவன:

$$w_{0\cdot10} = 0\cdot046$$
$$w_{0\cdot20} = 0\cdot062$$
$$w_{0\cdot30} = 0\cdot079$$
$$w_{0\cdot40} = 0\cdot097$$
$$w_{0\cdot50} = 0\cdot119$$
$$w_{0\cdot60} = 0\cdot147$$
$$w_{0\cdot70} = 0\cdot184$$
$$w_{0\cdot80} = 0\cdot241$$
$$w_{0\cdot90} = 0\cdot347$$
$$w_{0\cdot95} = 0\cdot461$$
$$w_{0\cdot99} = 0\cdot743$$
$$w_{0\cdot999} = 1\cdot168$$

α என்ற மிகைத்தன்மை மட்டத்தில் T_3-ன் ம $w_{1-\alpha}$ பகுதிமான மதிப்பைவிட அதிகமாக இருந்தால் H_0-ஐ மறுக்கிறோம்.

மாதிரி: $X_1 = 0\cdot621$, $X_2 = 0\cdot503$, $X_3 = 0\cdot203$, $X_4 = 0\cdot477$, $X_5 = 0\cdot710$, $X_6 = 0\cdot581$, $X_7 = 0\cdot329$, $X_8 = 0\cdot480$, $X_9 = 0\cdot554$, $X_{10} = 0\cdot382$ என்ற 10 ராண்டம் மாதிரிகள் கண்டறியப்பட்டன.

$F(x)$ ஆனது சீரான பரவல் சார்பு (uniform distribution function) என்ற சூனிய எடுகோளையும், சீரான பரவலைப் பெற்றிருக்கவில்லை என்ற மாற்று எடுகோளையும் கொள்வோம்.

இப்பொழுது க்ராமர்-வான் மைஸஸ் பொருத்தச் செம்மைச் சோதனைக்கான புள்ளியியல் அளவை,

$$T_3 = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right) \right\}^2 \text{ ஆகும்.}$$

கொடுக்கப்பட்ட விவரத்தை இப்பொழுது வரிசைப்படுத்தி அமைத்து $F^*\left(X^{(i)}\right)$ன் மதிப்புகளையும் $\left\{ F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right) \right\}$ மதிப்புகளையும் அட்டவணைப்படுத்துவோம்.

i	$X^{(i)}$	$F^*\left(X^{(i)}\right)$	$F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right)$	$\left\{F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right)\right\}^2$
1	0·203	0·203	0·153	0·023409
2	0·329	0·329	0·179	0·032041
3	0·382	0·382	0·132	0·017424
4	0·477	0·477	0·127	0·016129
5	0·480	0·480	0·030	0·000900
6	0·503	0·503	− 0·047	0·002209
7	0·554	0·554	− 0·096	0·009216
8	0·581	0·581	− 0·169	0·028561
9	0·621	0·621	− 0·229	0.052441
10	0·710	0·710	− 0·240	
				$\sum_{i=1}^{10} \left\{F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right)\right\}^2 = 0·239930$

$$\text{எனவே, } T_3 = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ F^*\left(X^{(i)}\right) - \left(\frac{2i-1}{2n}\right)\right\}^2$$

$$= \frac{1}{12\,(10)} + 0{\cdot}239930$$

$$= 0{\cdot}008333 + 0{\cdot}239930$$

$$= 0{\cdot}248263 \text{ ஆகும்.}$$

$\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் $w_{1-\alpha} = w_{0{\cdot}95}$-ன் பகுதிமான மதிப்பு $0{\cdot}461$ ஆகும். T_3-ன் கண்டறிந்த மதிப்பு $0{\cdot}248263$ ஆகும். இது $0{\cdot}461$-ஐ விடக் குறைவாக உள்ளதால் H_0-ஐ $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் ஏற்றுக் கொள்கிறோம்.

பயிற்சி 1

இரு வகைப்பட்ட முருங்கைக் காய்களின் உற்பத்தியில் $\frac{1}{3}$ பங்கு குட்டையாகவும், $\frac{2}{3}$ பங்கு நீளமாகவும் இருப்பதாக எதிர்பார்க்கப்படுகிறது. இவ் வெதிர்ப்பார்ப்பை அறிய நடத்திய சோதனையில் 240 காய்கள் குட்டையாகவும், 660 காய்கள் நீளமாகவும் இருப்பதாகக் கண்டுபிடிக்கப்பட்டது. நீளமான காய் வகைகளை வகுப்பு 1 என்றும், குட்டையான காய் வகைகளை வகுப்பு 2 என்றும் குறிப்பிட்டு $p = \frac{2}{3}$ எனக் கொண்டு, $H_0: \; p = \frac{2}{3}$, $H_1: \; p \neq \frac{2}{3}$ என்ற எடுகோள்களை ஈருறுப்புச் சோதனையைப் பயன்படுத்திச் சோதனை செய்க.

2. உதிரி உறுப்புகளை உற்பத்தி செய்யும் ஓர் இயந்திரம், 6 சதவீதமும் அதற்குக் குறைவான பழுதடைந்த உறுப்புகளை உற்பத்தி செய்தால், அது நல்ல நிலையில் வேலை செய்வதாகவும், 6 சதவீதத்திற்கு மேற்பட்ட உறுப்புகள் பழுதடைந்திருந்தால் அது பழுதுபார்க்கப்படவேண்டும் என்றும் கொள்ளப்படுகிறது. இயந்திரம் நல்ல முறையில் வேலை செய்கிறது என்ற சூனிய எடுகோளையும், இயந்திரம் பழுதுபார்க்கப்படவேண்டும் என்ற மாற்று எடுகோளையும் கொண்டு ஈருறுப்புச் சோதனையை மேற்கொள்க.

3. கல்வியின் தரம் குறித்து, மாநிலக் கல்விக் குழு அறிவித்த தகுதியை உயர்நிலைப் பள்ளிகள் பெற்றுள்ளனவா இல்லையா என்பதை அறிய 25 உயர்நிலைப் பள்ளிகள் ஒரு நகர்ப்புறத்திலிருந்து ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டன. இதில் 8 உயர்நிலைப் பள்ளிகள் மட்டுமே 'சிறந்தது' என்ற தகுதியைப் பெற்றுள்ளனவாக அறியப்பட்டது. நகரங்களிலுள்ள எல்லாப் பள்ளிகளையும் கருத்தில் கொள்ளும்போது இத் தகுதியை அப்

பள்ளிகள் பெற்றுள்ள p என்ற விகிதப் பள்ளிகளுக்கான 95% நம்பிக்கை இடைவெளியைக் காண்க.

4. X என்ற ராண்டம் மாதிரிக்கான 20 சார்பற்ற கண்டறிந்தவை $F(x)$ என்ற தெரியாத பரவல் சார்பைப் பெற்றுக் கீழ்க் கண்ட எண்களை அளித்துள்ளன.

142	134	98	119	131
103	154	122	93	137
86	119	161	144	158
165	81	117	128	103

$F(100)$-க்கான 95% நம்பிக்கை இடைவெளியைக் காண்க.

5. 42 விற்பனையாளர்களை அவர்களின் விற்பனைத் திறன்களுக்கு ஏற்ப மூன்று பிரிவுகளாகப் பிரிக்கப்பட்டு அவைகளுக்கான கண்டறிந்த அலைவெண்களும், எதிர்பார்க்கும் அலைவெண்களும் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டன.

	விற்பனைத் திறன்			மொத்தம்
	நல்ல முறையிலான	திருப்திகரமான	மோசமான	
கண்டறிந்த அலைவெண்கள்	18	12	10	40
எதிர்பார்க்கும் அலைவெண்கள்	16·5	12·5	11	40

இப் பிரிவுகளின் எதிர்பார்க்கும் அலைவெண்களும், கண்டறிந்த அலைவெண்களும் மிகைத்தன்மை செய்யுமளவிற்கு வேறுபட்டுள்ளனவா என்பதை கைவர்க்க ஒரு கூறுச் சோதனை மூலம் ஆய்க.

6. ஒரு பகடையை 300 தடவைகள் வீசி எறியும் சோதனையில் பகடை எண் 1, 2, 3, 4, 5, 6 ஆகியவை நிகழ்ந்துள்ள கண்டறிந்த அலைவெண்களும் அதற்கு ஏற்ற முறையான எதிர் பார்க்கும் அலைவெண்களும் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டுள்ளன.

பகடை எண்	கண்டறிந்த அலைவெண்கள்	எதிர்பார்க்கும் அலைவெண்கள்
1	40	48
2	56	52
3	39	30
4	58	56
5	63	68
6	44	46

கண்டறிந்த அலைவெண்களும், எதிர்பார்க்கும் அலைவெண்களும் ஒன்றுக்கொன்று மிகைத்தன்மை செய்கிறதா என்பதை கைவர்க்க ஒரு கூறுச் சோதனை மூலம் காண்க.

7. பெண்கள் பொது சேவையில் பங்கு கொள்ளலாம் என்பதை அறியும் முகமாக 25 சமூக சேவகர்களிடம் வினு எழுப்பியதில் அவர்கள் அளித்த மூன்று விதமான கருத்துகளின் விவரங்கள் பின் வருமாறு :

	கருத்துகள்		
	சேவை செய்யலாம்	சேவை செய்யக்கூடாது	சேவை தேவையற்றது
கண்டறிந்த சமூக சேவகர்களின் எண்ணிக்கை	11	8	6
எதிர்பார்க்கும் சமூக சேவகர்களின் எண்ணிக்கை	12·5	8·6	4·9

மேலே கொடுக்கப்பட்ட விவரங்கள் ஒரு குறிப்பிட்ட போக்கினைத் தெரிவிக்கின்றனவா என்பதை கைவர்க்க ஒரு கூறுச் சோதனை மூலம் ஆய்க.

8. தெளிவானது முதல், மங்கலானது வரையிலான புகைப் படங்கள் வரிசைக்கிரமத்தில் தரப்படுத்தப்பட்டன. கண்டறிந்த பரவலின் மதிப்புகளுக்கும், அறிமுறைக்குரிய பரவல் (theoretical distribution) மதிப்புகளுக்கும் இடையே வேறுபாடுகள் இல்லை என்ற சூனிய எடுகோளின் கீழ் அறிமுறைக்கான குவிவப் பரவலும் கண்டறிந்த விருப்பங்களுக்கு (observed choices) ஏற்ற குவிவப் பரவலும் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டன.

தரம்	1	2	3	4	5
f : தரத்தைப் பெற்ற நபர்களின் எண்ணிக்கை	0	1	1	3	5
$F^*(x) = H_0$-ன் கீழ் அமைந்துள்ள அறிமுறைக்குரிய குவிவப் பரவல்	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{5}{5}$

'கோல்மோகோராவ்-ஸ்மிர்னுவின்' ஒரு கூறுச் சோதனையைப் பயன்படுத்தி அறிமுறைக்குரிய பரவல் மதிப்புகளுக்கும் கண்டறிந்த பரவல் மதிப்புகளுக்கும் இடையே வேறுபாடுள்ளதா என்பதை ஆய்க.

9. ஒரு பரிசோதனையில் 25 மாணவர்களின் விளையாட்டுத் திறனைக் குறித்து ஆய்வு செய்யும் நோக்கத்தில், ஜோடியாக மாண வர்கள் விளையாடும் நிகழ்ச்சிகள் கண்காணிக்கப்பட்டன. இதில் பெரும்பாலான மாணவர்கள் ஒரே கல்லூரியிலிருந்து வந்துள்ளனர் என்று அறிந்த பின்பு, இந்தச் சோதனையின் ஆய்வு ஒரு சார்பான தாகவோ (biased) அல்லது ராண்டம் முறையில் அமையாமலோ இருக்கும் என்று கருதப்பட்டது. இதனை அறிய அம் மாணவர் களின் திறமைக்கான மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டுள்ளன.

மாணவர்கள்	மதிப்புகள்	மாணவர்கள்	மதிப்புகள்
1	52	14	26
2	44	15	17
3	57	16	18
4	55	17	20
5	66	18	22
6	58	19	86
7	24	20	13
8	38	21	14
9	54	22	26
10	25	23	29
11	38	24	89
12	97	25	36
13	71		

மேலுள்ள விவரங்களைக் கொண்டு ஒரு கூறு ஓட்டச் சோதனை யின் மூலம் குழுவின் இடைநிலை மதிப்பைவிட அதிகமாக அல்லது குறைவாக ஏற்படுவது, ராண்டம் முறையில் அமைகிறதா இல்லையா என்பதை ஆய்க.

10. ரயில் பயணச்சீட்டு கொடுக்கும் இடத்தில் ஆண்களும், பெண்களுமாக 60 பேர்கள் ஒரு வரிசையில் நின்றுகொண்டிருந்த விவரம் பின்வருமாறு :

ஆ, பெ, பெ, ஆ, ஆ, பெ, பெ, பெ, ஆ, ஆ, ஆ, பெ, ஆ, ஆ, ஆ, பெ, ஆ, பெ, ஆ, ஆ, ஆ, பெ, ஆ, ஆ, பெ, பெ, பெ, ஆ, ஆ, பெ, பெ, பெ, பெ, ஆ, ஆ, பெ, ஆ, பெ, ஆ, பெ, ஆ, ஆ, பெ, ஆ, பெ, ஆ, ஆ, பெ, பெ, ஆ, ஆ, ஆ, ஆ, ஆ, பெ, ஆ, ஆ, ஆ, பெ, பெ.

மேலுள்ள விவரங்களைக் கொண்டு, கொடுக்கப்பட்டுள்ள வரிசை, ராண்டம் மாதிரியில் அமைந்துள்ளதா இல்லையா என்பதை ஒரு கூறு ஓட்டச் சோதனை மூலம் ஆய்க.

11. தேர்வு முடிவுகளைக் கொண்ட செய்தித் தாளிலிருந்து 40 இரட்டைப்படை எண்கள் ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டன. இந்த எண்கள் உயர்நிலைப் பரவலைப் பெற்ற ராண்டம் மாறிகளிலிருந்து எடுக்கப்பட்ட கண்டறிந்தவைகளா என்பதை 'வில்லிகாக்ஸ்' சோதனை மூலம் கீழ்க்கண்ட விவரங்களை வரிசைப்படுத்தி அமைத்த பின்பு சோதனை செய்க.

75,	23,	23,	42,	56,	76,	59,	58,	68,	69,
72,	74,	25,	35,	92,	56,	79,	58,	45,	48,
35,	45,	36,	23,	29,	45,	73,	76,	63,	68,
58,	62,	47,	39,	41,	63,	68,	72,	97,	89.

12. $X_1 = 0{\cdot}635$, $X_2 = 0{\cdot}542$, $X_3 = 0{\cdot}214$,
$X_4 = 0{\cdot}463$, $X_5 = 0{\cdot}721$, $X_6 = 0{\cdot}591$,
$X_7 = 0{\cdot}326$, $X_8 = 0{\cdot}450$, $X_9 = 0{\cdot}556$,
$X_{10} = 0{\cdot}389$

என்ற 10 ராண்டம் மாதிரிகள் கண்டறியப்பட்டன. $F(x)$ என்ற பரவல் சார்பானது சீரான பரவல் சார்பைப் பெற்றுள்ளது என்ற சூனிய எடுகோளையும், சீரான பரவலைப் பெற்றிருக்கவில்லை என்ற மாற்று எடுகோளையும் கொண்டு 'க்ராமர்-வான் மைஸஸின்' சோதனையைப் பயன்படுத்தி ஆய்க.

13. $F(x)$ என்ற பரவல் சார்பிலிருந்து வருவிக்கப்பட்ட ராண்டம் மாதிரி $6{\cdot}3$, $4{\cdot}2$, $4{\cdot}7$, $6{\cdot}0$, $5{\cdot}7$ எனில் $F(x) = F^*(x)$ என்ற எடுகோளை (i) 'கோல்மோகோராவ்-ஸ்மிர்னுவ்' சோதனையின் வாயிலாகவும் (ii) 'க்ராமர்-வான் மைஸஸ்' சோதனையின் வாயிலாகவும் ஆய்க.

இதில் $F^*(x) = 0, \forall x < 4$

$$= \frac{x-4}{4}, \quad 4 \leqslant x < 8$$

$= 1, 8 \leqslant x$ ஆகும்.

2. இரு தொடர்புள்ள மாதிரிகளுக்கான சோதனைகள்

(Related Two Sample Tests)

இரு மாதிரிகளும் ஒன்றுக்கொன்று வேறுபட்டு அமைந்து உள்ளனவா அல்லது ஒரு மாதிரி மற்றொரு மாதிரியைவிட வேறு பட்டு அமைந்துள்ளதா என்பதனை அறிய இரு மாதிரிப் புள்ளி யியல் சோதனைகள் பயன்படுத்தப்படுகின்றன. மாதிரியானது ஒரு குழுவிலுள்ள நபர்களுக்கு அளிக்கப்பட்ட ஊசி மருந்தினால் ஏற்படும் விளைவு, அல்லது பயிற்சியின் விளைவெனக் கொள்ளலாம். ஒரு குழுவிலுள்ள நபர்கள் ஊசி மருந்தைப் பெறுதற்கு முன்பிருந்த நிலையை மாதிரி-1 என்றும் ஊசி மருந்து பெற்ற பிறகு ஏற்பட்டுள்ள மாற்று நிலையை மாதிரி-2 என்றும் கொண்டு, இவ் விரு மாதிரி களிடத்து ஏதாவது மிகைத்தன்மை (signify) செய்யுமளவிற்கு ஏதாவது வேறுபாடு உள்ளதா என்பதைத் தக்க சோதனைகளின் மூலம் அறியலாம்; அல்லது ஒரு குழுவிலுள்ள மாணவர்களுக்குப் பாடங்களை இரு முறையிலான வழிகளில் கற்பித்ததனால், கற்பித் தல் முறையில் ஏதாவது வேறுபாடு உள்ளதா இல்லையா என்பதனை அறியலாம். இவைகளை அறிய இவ் விரு மாதிரிகளையும் தொடர்பு படுத்தியோ, அல்லது மாதிரியின் ஒவ்வொரு முறையான மதிப்பை யும் ஜோடிப் பொருத்தம் (match) செய்தோ ஆய்வு செய்யலாம். இப்பொழுது இரு தொடர்புள்ள மாதிரிகளுக்குச் செய்யப்படும் சோதனைகளைக் காண்போம்.

மாறுதலில் மிகைத்தன்மைக்கான 'மெக்னமார்' சோதனை (McNemar Test for the Significance of Changes)

இச் சோதனை பெரும்பாலும் ஒரு மாதிரியிலுள்ள நபர் களின் முன்பிருந்த நிலைக்கும் (மாதிரி-1) இப்பொழுது இருக்கும் நிலைக்கும் (மாதிரி-2) ஏதாவது வேறுபாடு உள்ளதா இல்லையா என்பதை அறிவதற்குப் பயன்படுத்தப்படுகின்றது.

எடுத்துக்காட்டாக, ஒரு குழுவிலுள்ள எலிகளுக்கு ஒரு குறிப் பிட்ட ஊசி மருந்தைக் கொடுப்பதற்கு முன் இருந்த நிலையை

மாதிரி-1 என்றும், ஊசி மருந்து கொடுத்த பிறகு ஏற்பட்ட நிலையை மாதிரி-2 எனவும் கொண்டு பெறப்பட்ட இவ்விரு மாதிரிகளுக்கும் இடையே ஊசி மருந்தினால் மாற்றம் ஏதாவது உள்ளதா இல்லையா என்பதை அறியப் பயன்படுத்தப்படுகிறது.

இம் முறையில் கண்டறிந்தவைகளில் இருக்கக்கூடிய மாறுதல்களை அறிய நான்கு மடிப்பு அட்டவணையைப் (four fold table) பயன்படுத்துகிறோம்.

இவ்வட்டவணை, அதே நபர்களின் முதல், இரண்டாம் வகைப் பிரதி செயல்களுக்கான (responses) அலைவெண்களைப் பெற்று அமையும் இதனை +, − என்ற குறிகளின் மூலம் குறிப்பிட்டுக் கீழ்க் காணும் அட்டவணையை அமைப்போம்.

		பிறகு −	பிறகு +
முன்பு	+	a	b
	−	c	d

இதில் முதல், இரண்டாம் வகை பிரதிச் செயல்களுக்கிடையே உள்ள மாற்றங்கள் a, d ஆகிய அறைகளில் அமைந்துள்ளன. மாற்றமானது + லிருந்து − க்கு ஏற்பட்டால் அந்த நபரை அறை a-ல் அமைக்கிறோம். மாற்றமானது −லிருந்து +க்கு ஏற்பட்டால் அந்த நபரை அறை d-ல் அமைக்கிறோம். ஒரு மாறுதலும் இல்லையெனின், அஃதாவது + ஆனது + ஆகவே இருந்தால் அறை b-யிலும், − ஆனது − ஆகவேயிருந்தால் அறை c-யிலும் அமைக்கின்றோம்.

$a + d$ ஆனது மாறுதல் பெற்ற மொத்த நபர்களின் எண்ணிக்கை என்பதால் H_0-ன் கீழ் உள்ள எதிர்பார்ப்பான $\frac{a+d}{2}$, ஒரு திசையில் ஏற்பட்ட மாறுதலுக்கும் (அதாவது + லிருந்து − க்கும்), இதன் எதிர்த் திசையில் ஏற்பட்ட மாறுதலுக்கும் (அதாவது − லிருந்து + க்கும்) ஆன எதிர்பார்ப்பு $\frac{a+d}{2}$ என்றும் அமையும்.

அதாவது H_0-ன் கீழ் $\frac{(a+d)}{2}$ என்பது அறைகள் a, d ஆகியவைகளில் ஏற்படும். தெளிவாக சூனிய எடுகோள் $H_0 = +$ லிருந்து $-$ க்கான மாற்றத்திற்கான நிகழ்தகவு $P(+, -) = (-)$ லிருந்து $+$ க்கான மாற்றத்திற்கான நிகழ்தகவு $P(-, +)$ ஆகும்.

χ^2-ன் வரையறைப்படி,

$$\chi^2 = \sum_{k=1}^{l} \frac{(O_k - E_k)^2}{E_k}$$

k ஆம் வகைக் கண்டறிந்த மதிப்புகள் E_k என்பது H_0-ன் கீழ் k ஆம் வகை எதிர்பார்க்கும் மதிப்புகளாகும்.

மாறுதலின் மிகைத்தன்மைக்கான மெக்னமார் சோதனையில் அறைகள் a ஐயும், d ஐயும் கருத்தில் கொள்கிறோம்.

எனவே $a = a$ அறையிலுள்ள கண்டறிந்தவைகளின் எண்ணிக்கை

$d = d$ அறையிலுள்ள கண்டறிந்தவைகளின் எண்ணிக்கை எனில்,

$\frac{1}{2}(a + d) = a, d$ ஆகிய இரு அறைகளிலுமுள்ள எதிர் பார்க்கும் எண்ணிக்கைகளாகும். பிறகு,

$$\chi^2 = \sum_{a,d} \frac{(O-E)^2}{E}$$

$$= \frac{\left(a - \frac{a+d}{2}\right)^2}{\frac{a+d}{2}} + \frac{\left(d - \frac{a+d}{2}\right)^2}{\frac{a+d}{2}}$$

$$= \left\{ \frac{\left[\frac{a-d}{2}\right]^2}{\frac{a+d}{2}} \right\} + \left\{ \frac{\left[\frac{a-d}{2}\right]^2}{\frac{a+d}{2}} \right\} = \frac{(a-d)^2}{(a+d)}$$

இது 1 வரையற்ற பாகையைப் (degree of freedom) பெற்றமையும் χ^2 ஆகும்.

அதாவது H_0-ன் கீழ் மாதிரிப் பரவலானது, தோராயமான 1 வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாகக் கருதப்

படுவதால் தொடர்ச்சிக்கான திருத்தத்தை (correction for continuity) செய்தல் அவசியமாகின்றது. ஒரு தொடர்ச்சியற்ற பரவலுக்குத் தொடர்ச்சியான χ^2 பரவலைத் தோராயமாக்குவதால் இந்தத் திருத்தம் மிகவும் இன்றியமையாததாகிறது. எதிர்பார்க்கும் அலைவெண்கள் மிகவும் சிறியதாக இருந்தால் இந்த χ^2-ஐத் தோராயமாக்குவதால் பிழை ஏற்படுவது மிக நிச்சயமாகிறது. இப் பிழையை நீக்க இம் மாதிரியான தொடர்ச்சிக்கான திருத்தத்தைச் செய்கிறோம். இத் திருத்தத்தை யேட்ஸ் திருத்தம் (Yates correction) என்று கூறுகிறோம். திருத்தத்திற்குப் பிறகு, 1 வரையற்ற பாகையைக் (degree of freedom) கொண்ட

$$\chi^2 = \frac{(\,|a-d|-1)^2}{a+d} \tag{1}$$

சூத்திரம் (1) லிருந்து, χ^2-ன் கண்டறிந்த மதிப்பு, அட்டவணை C-ல் ஒரு வரையற்ற பாகைக்கு (degree of freedom), α என்ற மிகைத்தன்மை மட்டத்தின் மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால், முன்பு, பின்பு என்ற பிரதிச் செயல்களுக்கிடையே மிகைத்தன்மை (significance) உள்ளது என்று உய்த்துணருகிறோம்.

மாதிரி: 100 வாக்காளர்களைக் கொண்ட ராண்டம் மாதிரி ஒன்றில் ஒவ்வோர் உறுப்பினரும், கட்சி 1 அல்லது கட்சி 2 ஆகியவைகளை ஆதரிக்கிறார்களா என்ற செய்தி சேகரிக்கப்பட்டது. தேர்தல் பிரசாரத்திற்குப் பிறகு இந்த உறுப்பினர்கள் எந்தக் கட்சியை ஆதரிக்கிறார்கள் என்ற விவரங்கள் சேகரிக்கப்பட்டன. அவ் விவரங்கள் கீழ்க்கண்டவாறு அமைந்துள்ளன.

	பிரசாரத்திற்குப் பிறகு		
	கட்சி 1	கட்சி 2	மொத்தம்
பிரசாரத்திற்கு முன்பு — கட்சி 1	63	21	84
பிரசாரத்திற்கு முன்பு — கட்சி 2	4	12	16
மொத்தம்	67	33	100

'மெக்னமார்' சோதனையைப் பயன்படுத்தி, பிரசாரத்திஞல் ஓட்டு அளிக்கும் தன்மையில் ஏதாவது மாறுதல் ஏற்பட்டுள்ளதா என்பதனை ஆய்க.

சூனிய எடுகோளான H_0-ஐ, ஓட்டளித்த உறுப்பினர்கள் பிரசாரத்திஞல் மனமாற்றம் அடையவில்லை என்று கொள்வோம். மாற்று எடுகோளான H_1-ஐ, ஓட்டளித்த உறுப்பினர்களிடத்து மன மாறுதல் ஏற்பட்டு, கட்சி 1-ஐ ஆதரிக்கின்றனர் என்று கொள்வோம். $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். இங்கு $N = 100$,

$a = 21, \quad b = 63$

$c = 12, \quad d = 4$ ஆகும்.

எனவே, 'யேட்ஸின்' திருத்தத்தோடு, 1 வரையற்ற பாகையைக் (degree of freedom) கொண்ட

$$\chi^2 = \frac{(|a-d| - 1)^2}{a+d}$$

$$= \frac{(|21-4| - 1)^2}{21+4}$$

$$= \frac{(17-1)^2}{25}$$

$$= \frac{(16)^2}{25}$$

$$= \frac{256}{25}$$

$$= 10{\cdot}2 \text{ ஆகும்.}$$

அட்டவணை C-லிருந்து $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் 1 வரையற்ற பாகைக்கு $\chi^2 > 10{\cdot}2$ என H_0-ன் கீழ் நிகழ்வதற்கான நிகழ்தகவு $p < \frac{1}{2}({\cdot}001)$ ஆகும். அதாவது $p < {\cdot}0005$, H_0-ன் கீழ் நிகழ்ச்சி நிகழ்வதற்கான நிகழ்தகவு $0{\cdot}0005 < \alpha = 0{\cdot}05$ ஆகும். அல்லது $\chi^2_{1d \cdot f}$-ன் மதிப்பு $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் $3{\cdot}841$ ஆகும். கண்டறிந்த χ^2-ன் மதிப்பான $10{\cdot}2$ ஆனது $3{\cdot}841$-ஐ விட அதிகமாக உள்ளதால் H_0-ஐ மறுக்கிறோம். அதாவது, கண்டறிந்த χ^2-ன் மதிப்பு மறுக்கத்தகு வெளியில் அமைந்துள்ளது. எனவே, H_0-ஐ மறுக்கிறோம். அதனல் ஓட்டளிக்கும் மனப்பான்மை, பிரசாரத்திஞல் மாற்றம் பெற்றுள்ளது என உய்த்துணருகிறோம்.

குறியிடு சோதனை (Sign test)

$p = \frac{1}{2}, q = \frac{1}{2}$ என்று கொண்டு செய்யப்படும் ஈருறுப்புச் சோதனையைக் குறியிடு சோதனை எனலாம். குறியிடு சோதனையைப் பயன்படுத்தி இரு முழுமைத் தொகுதிகளும் (populations) ஒரே இடைநிலையைப் பெற்றுள்ளனவா என்பதை அறியலாம். ஒவ்வொரு முழுமைத் தொகுதியிலிருந்து ஜோடியாக விடுவிக்கப்பட்ட கண்டறிந்தவைகளைக் கொண்டு இச் சோதனை செய்யப்படுகின்றது, இச் சோதனையை எளிதாகக் கையாளலாம். நாம் ஆய்வு செய்கின்ற மாறியானது தொடர்ச்சியான பரவலைத் (continuous distribution) தழுவியதாக இருத்தல் வேண்டும் என்ற நிபந்தனையைத் தவிர வேறு எதையும் கருதத் தேவையில்லை. ஜோடியாகப் பிரிக்கத்தக்க வகையில் இரண்டு மாதிரிகளிலுள்ள விவரங்கள் அமையவேண்டும். சார்புடைய இரண்டு மாதிரிகளின் வேறுபாட்டின் மிகைத்தன்மையை அறிய இச் சோதனை பயன் படுத்தப்படுகிறது.

$(X_1, Y_1), (X_2, Y_2), \ldots\ldots\ldots (x_n, Y_n)$ என்ற ஒன்றை ஒன்று சார்பற்ற (mutually independent) இரு மாறி ராண்டம் மாதிரியைக் கருதுக. இதில் n' ஜோடி கண்டறிந்தவை உள்ளன. (X_i, Y_i) என்ற ஒவ்வொரு ஜோடியிலும் $X_i < Y_i$ எனில் $+$ குறி என்றும், $X_i > Y_i$ எனில் $-$குறி என்றும், $X_i = Y_i$ எனில் 0 அல்லது ஒத்தவை (ties) என்றும் கொள்கிறோம். மேலும் (X_i, Y_i) என்ற ஜோடியானது உள்ளே பொருத்தமுடையதாக (internally consistent) (X_i, Y_i) என்ற சில ஜோடிகளுக்கு $P(+) > P(-)$ என்று இருந்தால் எல்லா ஜோடிகளுக்கும் $P(+) > P(-)$ என்று இருக்கும் என்ற தற்கோளைக் (assumption) கொள்கிறோம். இதேபோன்று $P(+) < P(-)$-க்கும் $P(+) = P(-)$-க்கும் இம் மாதிரியான தற்கோளைக் கருதுகிறோம்.

(1) இரு முனைச் சோதனையைக் கருதினால் H_0 என்ற சூனிய எடுகோளும், H_1 என்ற மாற்று எடுகோளும் பின்வருமாறு அமைகின்றன.

$$H_0 : P(X_i < Y_i) = P(X_i > Y_i) \forall_i$$

$$H_1 : P(X_i < Y_i) < P(X_i > Y_i) \forall_i$$

அல்லது

$$P(X_i < Y_i) > P(X_i > Y_i) \forall_i \text{ என்றும்,}$$

(2) ஒருமுனைச் சோதனையைக் கருதினல் H_0 என்ற சூனிய எடுகோளும், H_1 என்ற மாற்று எடுகோளும் கீழ்க்கண்ட முறையில் இருவிதமாக அமைகின்றன.

$$H_0 : P(X_i < Y_i) \leqslant P(X_i > Y_i) \forall_i$$

$$H_1 ; P(X_i < Y_i) > P(X_i > Y_i) \forall_i$$ என்றும்,

(3) $$H_0 : P(X_i < Y_i) \geqslant P(X_i > Y_i) \forall_i$$

$$H_1 : P(X_i < Y_i) < P(X_i > Y_i \forall_i$$ என்றும் அமைகின்றன.

மேலுள்ள எடுகோள்களைச் சோதனை செய்யும்பொழுது குறியிடு சோதனையானது பிறழ்ச்சியற்றதாகவும் (unbiased) பொருத்தமுடையதாகவும் (consistent) இருப்பது கவனிக்கத் தக்கது. குறியிடு சோதனையானது மேலுள்ள எடுகோள்களுக்குத் தொடர்புடை எதிர்மறை (counter parts) எடுகோள்களைக் கருதும்பொழுது அது பிறழ்ச்சியற்றதாகவோ அல்லது பொருத்தமுடையதாகவோ இருக்காது. அத்தருணத்தில் (X_i, Y_i) களுக்கான பரவல்களைப் பற்றிக் கூடுதலான தற்கோளைக் (assumption) கொள்கிறோம். இருமுனைச் சோதனையைக் கருதினல் சூனிய எடுகோளை X_i-களும் Y_i களும் ஒரே இருப்பிடச் சுட்டுறுப்புகளைப் (location parameters) பெற்றிருப்பதாக விளக்கம் செய்கிறோம்.

எனவே, $H_0 : E(X_i) = E(Y_i) \forall_i$ ஐ அதற்கு எதிரான மாற்று எடுகோளை $H_i : E(X_i) \neq E(Y_i) \forall_i$ என்று கொண்டு X_i களும் Y_i களும் வேறு சராசரிகளைப் பெற்றுள்ளனவா என்பதைச் சோதனை செய்கிறோம். இதே போன்று இடைநிலைக்கான சோதனையாக $H_0 : X_i$ களின் இடைநிலை $= Y_i$ களின் இடைநிலை H_i : எல்லா i களுக்கும் x_i-களும், y_i-களும் வெவ்வேறு இடைநிலைகளைப் பெற்றுள்ளன என்ற எடுகோள்களைக் கொள்வதன் மூலம் சோதனை செய்யலாம். (1) லுள்ள ஒரு முனைச் சோதனையைக் கருதினல் சூனிய எடுகோளானது, X_i களின் மதிப்பு, Y_i களின் மதிப்பை விட மிகப் பெரியதாக உள்ளது என்று கருதலாம். ஏனெனில் H_0 ஆனது X_i-ன் மதிப்பானது Y_i-ன் மதிப்பை விட அதிகமாக இருக்கும் என்று வரையறுக்கப்படுவதால், இது ஒரு முனைக் குறியிடு சோதனையாக

$$H_0 : E(X_i) \geqslant E(Y_i) \forall_i$$

$$H_1 : E(X_i) < E(Y_i) \forall_i$$

என்று கருதி சோதனை செய்கிறோம். இதேபோன்று இடைநிலைக்கும் இம்மாதிரியான எடுகோள்களைக் கருதிச் சோதனை செய்யலாம்.

(2)-லுள்ள மற்றோர் ஒரு முனைச் சோதனையை, சூனிய எடு கோளானது x_i களின் மதிப்பு y_i களின் மதிப்பை விட மிகச் சிறியதாக இருக்கும் என்று கருதுவதன் மூலம்

$$H_0 : E(X_i) \leqslant E(Y_i) \ \forall_i$$
$$H_1 : E(X_i) > E(Y_i) \ \forall_i$$

என்று அமைக்கலாம். இதேபோன்று இடைநிலைக்கும் இம் மாதிரியான எடுகோள்களைக் கருதிச் சோதனை செய்யலாம்.

+ குறிகளைப் பெற்ற ஜோடிகளின் எண்ணிக்கையை T என்ற புள்ளியியல் அளவையாகக் கருதுவோம். அதாவது (X_i, Y_i) என்ற ஜோடிகளில் $X_i < Y_i$ என்று உள்ள ஜோடிகளின் எண்ணிக்கை யாகும். தெளிவாக $T = +$ குறிகளின் மொத்த எண்ணிக்கை யாகும். ஒத்தவைகளைப் (ties) பெற்ற ஜோடிகளைப் புறக்கணித் தால் மீதி ஜோடிகளின் எண்ணிக்கை $n = +$ குறிகளின் எண்ணிக் கையும் $-$ குறிகளின் எண்ணிக்கையும் சேர்ந்த மொத்த எண்ணிக்கையாகும்.

தேவைப்படும் மிகைத்தன்மை மட்டத்தை α என்று கொள்க. இரு முனைச் சோதனையைப் பயன்படுத்தவேண்டிய தருணத்தில் $n \leqslant 20$-க்குப் பொருத்தமான, அட்டவணை D-ஐ, $p = \frac{1}{2}$ என்று பார்த்து $\frac{\alpha}{2}$ க்கு ஏறக்குறைவான மதிப்பைக் கண்டு, அதனை α_1 என்று குறிப்போம். α_1-க்குப் பொருத்தமான y-ன் மதிப்பை t என்று குறிப்போம். $2\alpha_1$-ஐ அளவாகக்கொண்ட தீர்வுகட்ட மான வெளி $T \leqslant t$ அல்லது $T \geqslant n - t$ என்ற T-ன் மதிப்புகளைப் பெற்றதாக அமையும். $2\alpha_1$ என்ற மிகைத்தன்மை மட்டத்தில் $T \leqslant t$ அல்லது $T \geqslant n - t$ எனில் H_0-ஐ மறுக்கிறோம். மற்றபடி H_0-ஐ ஏற்றுக்கொள்ளுகிறோம். $n > 20$ எனில் அட்டவணை D-ன் கீழ் கொடுக்கப்பட்டுள்ள தோராயத்தைப் பயன்படுத்தி

$$t = \tfrac{1}{2}\left(n + w_{\frac{\alpha}{2}} \sqrt{n}\right) \qquad (1)$$

261688 BT-N (15) 31L6

ஐ-க் காண்கிறோம். இதிலுள்ள $w_{\frac{\alpha}{2}}$ ஐ, அட்டவணை T-லிருந்து பெறுகிறோம். $\alpha = 0{\cdot}05$ எனில் $w_{\frac{\alpha}{2}} = -1{\cdot}96$ ஆகும்.

எனவே, தோராயமாக $t = \dfrac{n}{2} - \sqrt{n}$ ஆகும். (2)

ஒரு முனைச் சோதனையை உய்த்துணரும்பொழுது, T-ன் பெரிய மதிப்புகள் $+$ குறிகள், $-$குறியைவிட அதிகமாக ஏற்படும் என்ற நிலையை H_1 உள்ளதுபோல் விளக்குகிறது. எனவே, T-ன் மதிப்புகளுக்கு ஒத்த தீர்வுகட்டமான வெளி $(n - t)$-க்குச் சமமாகவோ அல்லது அதிகமாகவோ இருக்கும். இதில் t-ஐ, அட்டவணை D லிருந்து $p = \frac{1}{2}$-க்கும், தகுந்த n-க்கும் பார்த்தால் பெறலாம். மேலும் α-க்கு சமீபத்தில் அமைந்த அட்டவணையின் மதிப்பையும் பெறலாம். இதனை α_1 என்று குறிப்போம். α_1-க்கு ஒத்த, y-ன் மதிப்பு T ஆகும். $n > 20$ எனில், t-ஐ $t = \frac{1}{2}\left(n + w_{\alpha}\sqrt{n}\right)$ என்று தோராயமாக அமைப்பதன் மூலம் பெறலாம். இதில் w_{α}-ஐ அட்டவணை T-லிருந்து பெறலாம். $T > n - t$ எனில், α_1 அல்லது α மிகைத்தன்மை மட்டத்தில் H_0-ஐ மறுக்கிறோம்.

$-$ குறிகளின் மொத்த எண்ணிக்கையை T' என்ற புள்ளியியல் அளவையைக் கருதி $T' \leqslant T$ எனில் H_0-ஐ மறுக்கிறோம். (2)-ல் உள்ளதால் $\alpha = 0{\cdot}025$ எனில் $t \cong \frac{n}{2} - \sqrt{n}$ ஐ எளிதில் காணலாம்.

(3) லுள்ள மற்றோர் ஒரு முனைச் சோதனையில் T-ன் சிறிய மதிப்புகள் $-$ குறியானது பெருமளவில் நிகழும் என்ற H_1-ன்படி இருப்பதை உணர்த்துகிறது. எனவே, t ஆனது மேலுள்ள முறைப்படி காணப்படுகின்றது. α_1 (அல்லது α) என்ற தீர்வு கட்டமான வெளியின் அளவுக்கு ஒத்த T-ன் மதிப்பு t-ன் மதிப்புக்குக் குறைவாகவோ அல்லது சமமாகவோ இருக்கும். α_1 என்ற மிகைத்தன்மை மட்டத்தில் $T \leqslant t$ எனில் H_0-ஐ மறுக்கிறோம்; அல்லது $n > 20$ எனில் α என்ற மிகைத்தன்மை மட்டத்தில் H_0-ஐ மறுக்கிறோம். மற்றபடி H_0-ஐ ஏற்றுக்கொள்கிறோம்.

மாதிரி : ஒரு குறிப்பிட்ட முறையைப் பயன்படுத்தி A என்ற பொருள் உற்பத்தி செய்யப்படுகின்றது. A செய்யும் அதே வேலையைச் செய்யக்கூடிய B என்ற பொருள் வேறு முறையைப் பயன்படுத்தி உற்பத்தி செய்யப்படுகின்றது. உற்பத்தி செய்பவர் A-க்கு B மேம்பட்டது என நுகர்வோர் கொள்வார்களா என்று அறியும் முகமாக 10 நுகர்வோர்களை ராண்டம் மாதிரியில் தேர்ந்தெடுத்து அவர்கள் ஒவ்வொருவருக்கும் A, B ஆகிய பொருள்கள் கொடுக்கப்பட்டுச் சில காலங்களுக்குப் பயன்படுத்துமாறு வேண்டப்பட்டனர். அவைகளுக்கான விருப்பங்களை அட்டவணைப் படுத்திய பின்பு, கண்ட $+$குறிகளின் எண்ணிக்கை 8, $-$குறிகளின்

எண்ணிக்கை 1, 'ஒத்தவை'களின் எண்ணிக்கை 1 என அறியப்பட்டது.

$$H_0 : P(+) \leqslant P(-)$$
$$H_1 : P(+) > P(-)$$

என்ற எடுகோள் கருதப்பட்டது.

இங்கு $n = +$ குறிகளின் எண்ணிக்கையும்,
$-$ குறிகளின் எண்ணிக்கையும் சேர்ந்ததாகும்.
$= 8 + 1 = 9.$
$T = +$ குறிகளின் எண்ணிக்கை $= 8$.

$n = 9, p = \frac{1}{2}$-க்கு அட்டவணை D-ல் $\alpha = 0 \cdot 05$-க்கு மிகச் சமீபமான மதிப்பு $\alpha_1 = 0 \cdot 0195$ என்று காண்கிறோம். இந்தத் தீர்வுகட்டமான வெளியின் அளவான $\alpha_1 = 0 \cdot 0195$, T-ன் மதிப்புகளுக்கு அதிகமான அல்லது சமமான மதிப்பான $n - t = 9 - 1 = 8$-க்கு ஒத்ததாகும்.

$T = 8$ என்பதால் H_0-ஐ மறுக்கிறோம். T-ன் கண்டறிந்த மதிப்பு மறுக்கத்தகு வெளியில் இருப்பதால் தீர்வு கட்டமான மட்டம்

$$\alpha = 0 \cdot 0195 \text{ ஆகும்.}$$

எனவே, நுகர்வோர்கள் B என்ற பொருளை A-க்கு மேம்பட்டது எனக் கொள்கின்றனர் என்று உற்பத்தி செய்பவர் தீர்மானிக்கிறார். இரு முனைக் குறியிடு சோதனைக்கான மாதிரியை இப்பொழுது காண்போம்.

மாதிரி: ஆர்பத்நாட் (1710) என்பவர் 82 வருடத்தின் பிறப்புப் பதிவேடுகளை எடுத்துக்கொண்டு ஒவ்வொரு வருடத்திலும் ஏற்பட்ட ஆண், பெண் குழந்தைகளின் எண்ணிக்கையை ஒப்பிட்டுப் பார்த்து, ஆண்களின் எண்ணிக்கை பெண்களின் எண்ணிக்கையை விட அதிகமாக இருந்தால் + என்றும் மாற்று விதமாக இருந்தால் − என்றும் குறித்தார். அவர் ஒப்பிட்டுப் பார்த்ததில் 'ஒத்தவை' (ties) எதுவுமில்லை. ஆர்பத்நாட் கண்ட + குறிகளின் எண்ணிக்கை 82 ஆகும்; − குறிகளே கிடையாது. இப்பொழுது கீழக்காணும் எடுகோள்களைக் காண்போம்.

$$H_0 : P(+) = P(-)$$
$$H_1 : P(+) \neq P(-)$$

புள்ளியியல் அளவையான $T = +$ குறிகளின் எண்ணிக்கை $= 82$
$\alpha = 0 \cdot 05$ என்ற தீர்வு கட்டமான வெளியின் அளவு,

T-ன் மதிப்புகள் $t = 0 \cdot 5 \left(82 - 1 \cdot 96 \sqrt{82}\right)$

$= 32 \cdot 1$-க்குக் குறைவாக இருப்பதற்கும்

T-ன் மதிப்புகள் $n - t = 82 - 32 \cdot 1$

$= 49 \cdot 9$-க்கு அதிகமாக இருப்பதற்கும்

ஒத்ததாகும். இங்குக் கண்டறிந்தது $T = 82$ ஆகும். எனவே H_0-ஐ மறுக்கிறோம்.

சிறப்பாக H_0-ஐ α-வின் சிறந்த அளவான

$$\widehat{\alpha} = P(T = 0) + P(T = 82)$$
$$= \left(\tfrac{1}{2}\right)^{82} + \left(\tfrac{1}{2}\right)^{82}$$
$$= \left(\tfrac{1}{2}\right)^{81}$$

என்ற மிகைத்தன்மை மட்டத்திலேயே மறுக்கப்பட்டிருக்கலாம்.

போக்குக்கான 'காக்ஸ்–ஸ்டுவர்டின்' சோதனை (Cox and Stuart test for trend)

தொடர்ச்சியாக அமைந்துள்ள $X_1, X_2, \ldots\ldots X_{n'}$ என்ற ராண்டம் மாறிகளைக் கொண்ட விவரங்கள் ஒரு குறிப்பிட்ட வரிசையில் கண்டறிந்த ராண்டம் மாறிகள் அமைந்த வரிசையிலேயே கருதப்படுகின்றன. இந்தத் தொடர்ச்சியில் ஏதாவது போக்கு (trend) இருக்கிறதா என்பதை அறிவதே இச் சோதனையின் நோக்கமாகும். ராண்டம் மாறிகளை ஜோடிகளாக $X_1, X_{1+k}), (X_2, X_{2+k}), \ldots\ldots\ldots (X_{n'-k}, X_{n'})$ என்று அமைக்கிறோம். இதில் n' இரட்டைப்படையாக இருந்தால் $k = \frac{n'}{2}$ என்றும், n' ஒற்றைப்படையாக இருந்தால் $k = \frac{n'+1}{2}$ என்றும் கொள்கிறோம். குறிப்பாக இதில் n' ஒற்றைப்படையாக இருந்தால் நடு ராண்டம் மாறியை நீக்கிவிட்ட பிறகுதான் மேலுள்ள ஜோடி முறையை அமைக்கிறோம்.

(X_i, X_{i+k}) என்ற ஜோடிக்கு $X_i < X_{i+k}$ என்று இருந்தால் $+$ குறியும் $X_i > X_{i+k}$ என்றிருந்தால் $-$ குறியும் இடுகிறோம். 'ஒத்தவை' (ties) இருந்தால் அவைகளை நீக்கிவிடுகிறோம். எஞ்சியுள்ள 'ஒத்தவை' இல்லாத ஜோடிகளின் எண்ணிக்கையை n என்று குறிக்கிறோம்.

இச் சோதனையில் $X_1, X_2, X_3, \ldots\ldots, X_n'$ என்ற ராண்டம் **மாறிகள்** ஒன்றுக்கொன்று சார்பற்றவை என்றும், X_i-களின் **அளவுகோல்** குறைந்தது வரிசையில் (ordinal) அமைந்ததாகவும், X_i-கள் சர்வ சமமான பரவலைப் (identically distributed) பெற்றுள்ளதாகவோ அல்லது ஒரு போக்கு (trend) இருப்பதாகவோ தற்கோளைக் கொள்கிறோம். இப்பொழுது கீழ்க்காணும் எடுகோள் குறியிடு சோதனையிலுள்ள எடுகோள்களுக்குத் தொடர்புடை எதிர்மறை (counter parts) எடுகோள்களாகக் கருதலாம்.

(1) இரு முனைச் சோதனைக்கு

$$H_0 : P(X_i < X_{i+k}) = P(X_i > X_{i+k}) \forall_i$$

$$H_1 : P(X_i < X_{i+k}) \neq P(X_i > X_{i+k}) \forall_i$$

அதாவது, H_0 : போக்குகள் நிலைத்திருக்கவில்லை.

H_1 : மேல் போக்கோ (upward trend) அல்லது கீழ்ப் போக்கோ உள்ளது.

2) ஒரு முனைச் சோதனையில்

$$H_0 : P(X_i < X_{i+k}) \leqslant P(X_i > X_{i+k}) \forall_i$$

$$H_1 : P(X_i < X_{i+k}) > P(X_i > X_{i+k}) \forall_i$$

அதாவது, H_0 : மேல் போக்கு எதுவுமில்லை,

H_1 : மேல் போக்கு உள்ளது.

(3) ஒரு முனைச் சோதனையில்

$$H_0 : P(X_i < X_{i+k}) \geqslant P(X_i > X_{i+k}) \forall_i$$

$$H_1 : P(X_i < X_{i+k}) < P(X_i > X_{i+k}) \forall_i$$

அதாவது, H_0 : கீழ்ப் போக்கு எதுவுமில்லை,

H_1 : கீழ்ப் போக்கு உள்ளது

என்றும் கருதுகிறோம்.

குறியிடு சோதனையில் உள்ளதுபோல் இங்கும் புள்ளியியல் அளவுச் சோதனையான T என்பது + குறிகளின் மொத்த எண்ணிக்கையாகும். அதாவது ஜோடிகளில் $X_{i+k} > X_i$ களின் மொத்த எண்ணிக்கையாகும். குறியிடு சோதனையிலுள்ள தீர்மான விதியே காக்ஸ்-ஸ்டூவர்ட் சோதனைக்கும் பயன் படுத்தப்படுகிறது. அதாவது T-ஐக் கண்ட குறியிடு சோதனையிலுள்ள தீர்மான விதியே இச் சோதனைக்கும் பயன்படுத்தப் படுகிறது.

மாதிரி: ஒரு வருடத்தில் ஏற்படும் மொத்த மழை அளவுகள் அங்குலங்களில் ஒவ்வொரு வருடத்திற்குமாக, 19 வருடங்களுக்குக் கீழ்க்கண்டவாறு பதிவு செய்யப்பட்டன. இந்த மழை பெய்தலின் போக்கு வருடம் அதிகமாக ஆக உயருகிறதா அல்லது குறைகிறதா என்பதை அறிய 'காக்ஸ்-ஸ்டுவர்ட்' சோதனையைப் பயன்படுத்துவோம்.

45·25, 45·83, 41·77, 36·26, 45·37, 52·25,
35·57, 57·16, 35·37, 58·32, 41·05, 33·72,
45·73, 37·90, 41·72, 36·07, 49·83, 36·24,
39·90.

இங்கு $n' = 19$ ஆகும். இது ஒற்றைப்படை எண் என்பதால் வரிசையின் நடு எண்ணை 58·32-ஐ நீக்கிவிட்டு எஞ்சியுள்ள எண்களை ஜோடிகளாகப் பின்வருமாறு அமைக்கிறோம்.

(45·25, 41·05)	(45·27, 41·72)
(45·83, 33·72)	(52·25, 36·07)
(41·77, 45·73)	(35·37, 49·83)
(36·26, 37·90)	(57·16, 36·24)
	(35·37, 39·90)

இங்கு 'ஒத்தவை' (ties) கிடையாது. எனவே $n = 19$ ஆகும். புள்ளியியல் அளவான T = இரண்டாம் எண், முதல் எண்ணைவிட அதிகமாக இருக்கும் ஜோடிகளின் எண்ணிக்கை = 4 ஆகும். தீர்வு கட்டமான வெளியின் அளவான 0·0390 ஆனது

T-ன் மதிப்பு $\leqslant 1$-க்கும்

$T \geqslant 9 - 1 = 8$-க்கும் ஒத்ததாகும்.

$T = 4$ என்ற கண்டறிந்த மதிப்பானது ஏற்கத்தகு வெளியில் இருக்கிறது. மேலும் தீர்வு கட்டமான மட்டம் 1·0 ஆகும்.

எனவே, 'போக்கு நிலைத்திருக்கவில்லை' என்ற சூனிய எடுகோளை ஏற்றுக்கொள்கிறோம்.

மாதிரி: ஒவ்வொரு மாதமும் அணைக்கட்டிலிருந்து ஒரு வினாடிக்கு விடப்படும் தண்ணீர் வீதம் கன அடியில், 24 மாதங்களுக்குப் பின்வருமாறு குறிக்கப்பட்டது.

மாதம்	முதல் வருடம்	இரண்டாம் வருடம்
ஜனவரி	14·6	14·2
பிப்ரவரி	12·2	10·5
மார்ச்	104·0	123·0
ஏப்ரல்	220·0	190·0
மே	110·0	138·0
ஜூன்	86·0	98·1
ஜூலை	92·8	88·1
ஆகஸ்டு	74·4	80·0
செப்டம்பர்	75·4	75·6
அக்டோபர்	51·7	48·8
நவம்பர்	29·3	27·1
டிசம்பர்	16·0	15·7

'காக்ஸ்-ஸ்டூவர்டின்' போக்குச் சோதனையைப் பயன் படுத்திக் கீழ்க்காணும் எடுகோள்களை இப்பொழுது சோதனை செய்வோம்.

H_0 : தண்ணீர் விடப்படும் வீதம் (rate) குறையவில்லை.

H_1 : தண்ணீர் விடப்படும் வீதம் குறைந்துள்ளது.

இங்குப் புள்ளியியல் அளவு T = இரண்டாம் வருடத்தில் விடப் படும் தண்ணீர் வீதம், முதலாம் வருடத்தில் விடப்பட்ட தண்ணீர் வீதத்தைவிட அதிகமாகப் பெற்றுள்ள ஜோடிகள் = 5 ஆகும். கீழ்ப்போக்கு உள்ளதா என்பதைச் சோதனை செய்வதே இச் சோதனையின் நோக்கம் என்பதால் தீர்வு கட்டமான வெளியின் மதிப்பான 0·0730 ஆனது T-ன் மதிப்புகள் 3-க்குச் சமமாக அல்லது குறைவாக இருப்பதற்கு ஒத்திருக்கிறது. இதனை அட்ட வணை D-ல் $n = 12$, $p = \frac{1}{2}$-க்குப் பார்த்தால் பெறலாம். எனவே, H_0-ஐ ஏற்றுக் கொள்கிறோம்.

தீர்வு கட்டமான மட்டம்

$$\widehat{\alpha} = P(T \leqslant 5 \mid H_0 \text{ உண்மையானால்})$$

$= 0·3872$ ஆகும். இது ஏற்றுக்கொள்ளக் கூடிய α-ஐ விடப் பெரியதாக உள்ளது.

மாதிரி: இருவிதமான மருந்துகள் 10 நோயாளிகளுக்கு அளிக்கப்பட்டு அதனுல் ஏற்பட்ட விளைவுகளுக்கு (reactions) நேர் ஒட்டுறவு உள்ளதா என்பதைக் கீழ்க்காணும் விவரங்களைக் கொண்டு 'காக்ஸ்-ஸ்டுவர்ட்' சோதனை மூலம் அறிவோம்.

நோயாளி	மருந்து 1	மருந்து 2
1	+ 0·7	+ 1·9
2	— 1·6	+ 0·8
3	— 0·2	+ 1·1
4	— 1·2	+ 0·1
5	— 0·1	— 0·1
6	+ 3·4	+ 4·4
7	+ 3·7	+ 5·5
8	+ 0·8	+ 1·6
9	·0	+ 4·6
10	+ 2·0	+ 3·4

கொடுக்கப்பட்ட விவரங்களை மருந்து 1-க்கு ஏறும் வரிசையில் வரிசையாக அதற்கேற்ற மருந்து 2-ன் மதிப்புகளோடு அமைப்போம்.

நோயாளி	மருந்து 1	மருந்து 2
2	— 1·6	+ 0·8
4	— 1·2	+ 0·1
3	— 0·2	+ 1·1
5	— 0·1	— 0·1
9	·0	+ 4·6
1	+ 0·7	+ 1·9
8	+ 0·8	+ 1·6
10	+ 2·0	+ 3·4
6	+ 3·4	+ 4·4
7	+ 3·7	+ 5·5

இப்பொழுது புதிதாக வரிசைப்படுத்தி அமைக்கப்பட்ட மருந்து 1-க்கு ஏற்ற மருந்து 2-ன் மதிப்புகளை ஜோடியாக அமைத்தால்

(+ 0·8, + 1·9) (+ 0·1, + 1·6)
(+ 1·1, + 3·4) (— 0·1, + 4·4)
(+ 4·6, + 5·5) என்று பெறுகிறேம்.

இதற்கு ஒரு முனை 'காக்ஸ்-ஸ்டுவர்ட்' சோதனையைப் பயன் படுத்துவோம். அதாவது,

H_0 : நேர் ஒட்டுறவு இல்லை,

H_1 : நேர் ஒட்டுறவு உள்ளது

என்று கொள்கிறோம். அதாவது மேல்போக்கு இருக்கிறதா என்பதைச் சோதனை செய்கிறோம். இங்கு $T = 5$ ஆகும். ஏனெனில், எல்லா 5 ஜோடிகளிலும் இரண்டாம் மதிப்பு முதல் மதிப்பைவிட அதிகமாக உள்ளது.

தீர்வுகட்டமான வெளியின் அளவான $0{\cdot}0312$ (அட்டவணை 3-ல் $n = 5$, $p = \frac{1}{2}$-க்குப் பார்த்தால் பெறலாம். $t = 0$ ஆகும்) T-க்கு ஒத்ததாகும். எனவே, சூனிய எடுகோளை மறுக்கிறோம். இதனால் மருந்து 1, மருந்து 2 ஆகியவைகளினால் ஏற்படக்கூடிய பிரதிகிரியைகளுக்குள் நேர் ஒட்டுறவு உள்ளது என்று தீர்மானிக்கிறோம். இங்கு, தீர்வுகட்டமான மட்டம் $\widehat{\alpha} = 0{\cdot}0312$ ஆகும்.

மாதிரி : 24 மணி நேரத்தில், ஒவ்வொரு மணி நேரத்திலும், கொசுக்கள் இட்ட முட்டைகளின் எண்ணிக்கை, கீழ்க்காணும் அட்டவணையில் கொடுக்கப்பட்டுள்ளது.

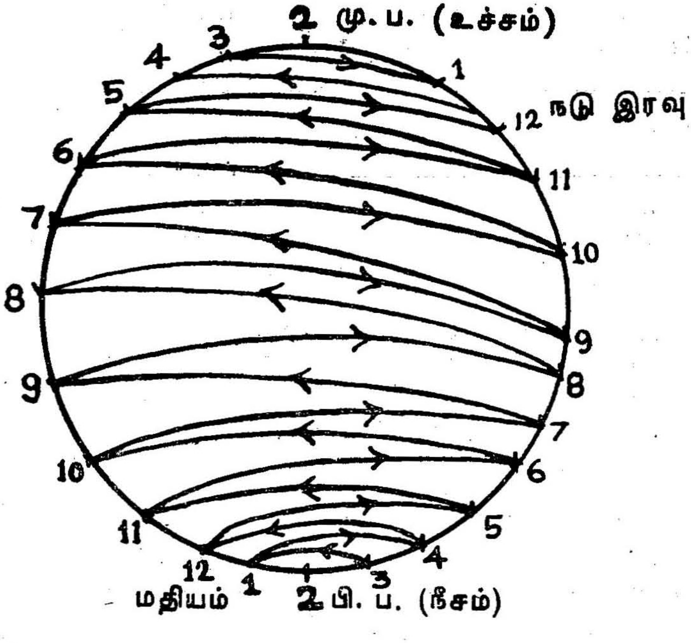
நேரம்	முட்டைகளின் எண்ணிக்கை	நேரம்	முட்டைகளின் எண்ணிக்கை
9 மு.ப	141	9 பி.ப	153
10 மு.ப	109	10 பி.ப	198
11 மு.ப	136	11 பி.ப	273
மதியம்	101	நள்ளிரவு	286
1 பி.ப	53	1 மு.ப	276
2 பி.ப	74	2 மு.ப	225
3 பி.ப	50	3 மு.ப	213
4 பி.ப	99	4 மு.ப	166
5 பி.ப	73	5 மு.ப	166
6 பி.ப	156	6 மு.ப	164
7 பி.ப	133	7 மு.ப	129
8 பி.ப	106	8 மு.ப	127

மேலுள்ள விவரங்களிலிருந்து, 'காக்ஸ்-ஸ்டுவர்டின்' போக்குக்கான சோதனையைப் பயன்படுத்திப் பின்வரும் எடுகோள்களைச் சோதனை செய்வோம்.

சூனிய எடுகோள் H_0 : 24 மணிகளுக்கான முட்டைகளின் எண்ணிக்கையானது ஒத்த பரவல்களைப் பெற்ற (identically distributed) 24 ராண்டம் மாறிகள் என்க.

மாற்று எடுகோள் H_1 : இடப்படும் முட்டைகளின் எண்ணிக்கை 2-15 பிற்பகலுக்கு நீசத்தைப் (minimum) பெற்று, சிறிது சிறிதாக உயர்ந்துகொண்டே சென்று 2-15 முற்பகலுக்கு உச்சத்தைப் பெற்று, மறுபடியும் 2-15 பிற்பகல்வரை குறைந்து கொண்டே இருக்கும் என்று கொள்ளப்படுகிறது.

மாற்று எடுகோள் உண்மையானது என்றால் 2-15 பிற்பகலுக்கு அருகில் இடப்படும் முட்டைகளின் எண்ணிக்கை மிகக் குறைந்தும், 2-15 முற்பகலுக்கு அருகில் இடப்படும் முட்டைகளின் எண்ணிக்கை மிக அதிகரித்தும் இருக்கும். எனவே, இப்பொழுது முட்டைகளின் எண்ணிக்கையை நேரத்திற்கேற்றாற் போல 2-15 பிற்பகலுக்கு அருகிலும், 2-15 முற்பகலுக்கு அருகிலும் இருக்குமாறு மாற்றி வரிசைப்படுத்தி அமைப்போம்.



அதாவது மேலுள்ள படத்திலுள்ள அமைப்பைப்போல் பின்வருமாறு வரிசைப்படுத்துகிறோம்.

காலம்	முட்டைகளின் எண்ணிக்கை	காலம்	முட்டைகளின் எண்ணிக்கை
2 பி.ப	74	8 மு.ப	127
3 பி.ப	50	9 பி.ப	153
1 பி.ப	53	7 மு.ப	129
4 பி.ப	99	10 பி.ப	198
மதியம்	101	6 மு.ப	164
5 பி.ப	73	11 பி.ப	273
11 மு.ப	136	5 மு.ப	166
6 பி.ப	156	நடு இரவு	286
10 மு.ப	109	4 மு.ப	166
7 பி.ப	133	1 மு.ப	276
9 மு.ப	141	3 மு.ப	213
8 பி.ப	106	2 மு.ப	225

மாற்று எடுகோளான H_1 : உண்மையானல் இந்த எண்கள் ஏற்றப் போக்கைக் காட்டவேண்டும். இங்கு 'காக்ஸ்-ஸ்டுவர்டின்' ஒரு முனைச் சோதனையைப் பயன்படுத்துகிறோம்.

மேலுள்ள அட்டவணையிலுள்ள முதல் நிரல் (column) எண்களை இரண்டாம் நிரலிலுள்ள எண்களுடன் ஜோடி செய்து இரண்டாம் நிரலிலுள்ள ஒவ்வோர் எண்ணும் அதற்கேற்ற முதல் நிரலிலுள்ள ஒவ்வோர் எண்ணைவிட அதிகமாக இருக்கும் எண்களின் எண்ணிக்கையைக் கணக்கிடுகிறோம்.

அதாவது T-ஐக் கீழ்க்கண்டவாறு கணக்கிடுகிறோம்.

127	>	74	+
153	>	50	+
129	>	53	+
198	>	99	+
164	>	101	+
273	>	73	+
166	>	136	+
286	>	156	+
166	>	109	+
276	>	133	+
213	>	141	+
225	>	106	+

இங்கு $T = 12$; $n = 12$; $p = \frac{1}{2}$.

அட்டவணை D-லிருந்து தீர்வுகட்டமான வெளியின் அளவான $\alpha = 0{\cdot}0193$ ஆனது $T \geqslant 12-2 = 10$-க்கு ஒத்த மதிப்பென்று அறிகிறோம்.

எனவே, H_0-ஐ மறுக்கிறோம். எனவே, இடப்படும் முட்டைகளின் எண்ணிக்கை 2-15 பிற்பகலின்போது நீசத்தை நெருங்கியும், 2-15 முற்பகலின்போது உச்சத்தை நெருங்கியும் இருக்கிறது என்பது சரியானதே என்று, மாற்று எடுகோளை ஏற்றுக்கொள்கிறோம்.

தீர்வு கட்டமான மட்டம் $\widehat{\alpha} = P\,(T \geqslant 12) = 0{\cdot}0002$ ஆகும்.

'வில்காக்ஸனின்' ஜோடிப் பொருத்தக் குறியீடு தரச் சோதனை (Wilcoxon's Matched pair - singed - rank test)

X_i, Y_i என்ற இரு நடத்துகைகளைக் (treatment) கருத்தில் கொள்ளும்பொழுது, (X_i, Y_i) என்ற ஒவ்வொரு ஜோடியின் மதிப்புகளுக் கிடையேயுள்ள வித்தியாசத்தை $X-Y_i = d_i$ என்று குறிப்பிடுவோம். வில்காக்ஸன் சோதனையைப் பயன்படுத்த $X_i-Y_i = d_i$ என்ற வித்தியாசங்களைக் குறிகளைப் பொருட்படுத்தாமல் வரிசைப் படுத்துகிறோம். பின்பு இந்த வரிசைப்படுத்திய மதிப்புகளில் மிகச்சிறிய மதிப்பிற்கு 1 என்ற தரத்தையும் (rank), அதற்கடுத்த பெரிய மதிப்பிற்கு 2 என்ற தரத்தையும், இதுபோலவே மேன்மேலும் உயர மதிப்புகளுக்கான தரங்களைக் கொடுக்கிறோம். பிறகு ஜோடிகளின் மதிப்புகளுக்கிடையேயுள்ள வித்தியாசத்தின் குறிகளை இந்தத் தரங்களுக்குப் பொருத்துகிறோம்.

இப்பொழுது *A*, *B* என்ற இரு நடத்துகைகளும் சமமானவை என்றால், அதாவது H_0 என்ற சூனிய எடுகோள் உண்மையானால் 'd_i' களில் மிகப்பெரிய நடத்துகை *A*-ஐத் தழுவியும் சிறிய நடத்துகை *B*-ஐத் தழுவியும் இருக்கும் என்று எண்ணுகிறோம். அதாவது பெரிய தரங்களில் சில, நேர் d_i களிலிருந்து வரப்பெற்றவை என்றும், மற்றவை எதிர் d_i களிலிருந்து வரப்பெற்றவை என்றும் கருதுகிறோம். இவ்வாறாக நேர் தரங்களின் (positive ranks) கூட்டலையும், எதிர் தரங்களின் (negative ranks) கூட்டலையும் கணக்கிட்டால், இரு கூடுதல்களின் மதிப்பும் H_0-ன் கீழ் சமமாக இருக்கும். ஆனால், நேர் தரங்களின் கூட்டல் எதிர் தரங்களின் கூட்டலினின்று வேறுபட்டு இருந்தால், நடத்துகை *A* ஆனது, நடத்துகை *B*-லிருந்து மிகவும் வேறுபட்டுள்ளது என உய்த்துணருகிறோம். அதாவது எதிர் d_i களுக்கான

தரங்களின் கூட்டலோ அல்லது நேர் d_i களுக்கான தரங்களின் கூட்டலோ மிகச் சிறியதாக இருந்தால் H_0-ஐ மறுக்கிறோம்.

சில சந்தர்ப்பங்களில் ஒத்த ஜோடியின் இரு மதிப்புகளும் சமமாக இருக்கலாம். அச் சமயங்களில் $d = 0$ என்றிருக்கும். இச் சந்தர்ப்பங்களை 'ஒத்தவை' (ties) என்று கூறுகிறோம்.

ஒத்தவைகளைப் பெற்ற அம்மாதிரியான ஜோடிகளை நீக்கி விடுகிறோம். மொத்த ஜோடிப் பொருத்தங்களின் (matched pairs) எண்ணிக்கையிலிருந்து, $d = 0$ ஐப் பெற்ற ஜோடிகளின் எண்ணிக்கையைக் கழித்துக் கிடைக்கும் ஜோடிகளின் எண்ணிக்கையை N என்று குறிப்பிடுவோம்.

மற்றொரு மாதிரியான ஒத்தவை (ties) இரண்டு அல்லது அதற்கு மேற்பட்ட d-க்கள் ஒரே மதிப்புகளைப் பெற்றிருந்தால் நிகழும். இம்மாதிரியான ஒத்தவைகளைப் பெற்ற ஜோடிகளுக்குச் சமமான தரங்களை அளிக்கிறோம்.

உதாரணமாக மூன்று ஜோடிகளின் d-ன் மதிப்புகள் -1, $+1$, $+1$ என்றிருந்தால் இவைகளின் ஒவ்வொரு ஜோடிக்கும் $\frac{1+2+3}{3} = 2$ என்ற தரத்தைச் சமமாகப் பங்கிட்டு அளிக்கிறோம். 1, 2, 3 என்ற தரங்களைச் சமமாகப் பங்கிட்டுக் கொடுத்துவிட்டதால் வரிசையில் அடுத்து நிகழும் d-க்கு 4 என்ற தரத்தை அளிக்கிறோம். சிறு கூறுகளுக்கு (small samples) நேர் அல்லது எதிர் தரங்களின் கூட்டலில் எந்தக் கூட்டல் தொகை சிறியதாக உள்ளதோ அக் கூட்டல் தொகையை T என்று குறிப்பிடுகிறோம். இதில் நேர் தரங்களின் கூட்டல், எதிர் தரங்களின் கூட்டலைவிடச் சிறியதாகவோ அல்லது எதிர் தரங்களின் கூட்டல், நேர் தரங்களின் கூட்டலைவிடச் சிறியதாகவோ இருக்கலாம். இப்பொழுது அட்டவணை G-ல் T-ன் பல்வேறு மதிப்புகளுக்கு அதற்கியைந்த மிகைத் தன்மை மட்டங்கள் கொடுக்கப்பட்டுள்ளன. α என்ற குறிப்பிட்ட மிகைத்தன்மை மட்டத்தில் கண்டறிந்த T-ன் மதிப்புகளுக்கேற்ற, அட்டவணை G-லுள்ள மதிப்பிற்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால் சூனிய எடுகோளை அம் மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

அட்டவணை G-ஐ ஒரு முனைச் சோதனைக்கும் இரு முனைச் சோதனைக்கும் பயன்படுத்தலாம்.

கொடுக்கப்பட்ட விவரங்களைப் பரிசோதனை செய்வதற்கு முன்பு பரிசோதனை செய்பவர் குறைந்த தரங்களின் கூட்டலின் குறியை ஊகமாக அறிந்தால் ஒரு முனைச் சோதனையைப் பயன் படுத்தலாம்.

உதாரணமாக, எதிர் தரங்களின் கூட்டல் $T = 3$ என்க. இது நேர் தரங்களின் கூட்டலைவிடச் சிறியது என்று ஊகமாக அறியப் பட்டது. $N = 9$ எனில், இரு குழுவிற்குமிடையே வேற்றுமை இருக்கும் என்ற மாற்று எடுகோளைக் கொண்டால், சூனிய எடுகோளை $\alpha = 0{\cdot}02$ மிகைத்தன்மை மட்டத்தில் மறுப்போம். மேலும் $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தில் எதிர் தரங் களின் கூட்டல்தான் சிறியது என்ற மாற்று எடுகோளைக் கொண் டால் சூனிய எடுகோளை மறுப்போம்.

மாதிரி: இரட்டையர்களில் முதலில் பிறந்தவர், இரண்டாவ தாகப் பிறந்தவரைவிட அதிக ஆக்கிரமிப்புத் தன்மை பெற்றுள் ளாரா? என்று அறியும் முகமாக 8 இரட்டைப் பிறவிகளுக்கு உளவியல் சோதனை கொடுக்கப்பட்டு, ஆக்கிரமிப்புக்கான அளவுகள் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டன.

வரிசை	முதலில் பிறந்தவர்	இரண்டாவதாகப் பிறந்தவர்
1	82	63
2	69	42
3	73	74
4	43	37
5	58	51
6	56	43
7	76	80
8	65	82

இப்பொழுது மேலுள்ள விவரங்களுக்கு வில்காக்ஸின் சோதனையைப் பயன்படுத்துவோம்.

H_0 என்ற ஆக்கிரமிப்புத் தன்மையில் இரட்டையர்களுக்கிடையே வேற்றுமையில்லை என்ற சூனிய எடுகோளையும், அவர்களுக்கிடையே ஆக்கிரமிப்புத் தன்மை குறித்து வேற்றுமைகள் உள்ளன என்ற மாற்று எடுகோளையும் கொள்வோம்.

அதாவது சூனிய எடுகோள்,

H_0 : நேர் தரங்களின் கூட்டல் = எதிர் தரங்களின் கூட்டல்

H_1 : நேர் தரங்களின் கூட்டல் $\neq$ எதிர் தரங்களின் கூட்டல் என்று கொள்வோம்.

இப்பொழுது *d*-க்களையும் அவைகளுக்கான குறியிட்ட தரத்தையும் கொண்ட அட்டவணையை அமைப்போம்.

வரிசை	முதலில் பிறந்தவர்	இரண்டாவதாகப் பிறந்தவர்	*d*	குறிகளின்றி *d*-க்கான தரம்	குறிகளுடன் *d*-க்கான தரம்
1	82	63	19	7	7
2	69	42	27	8	8
3	73	74	−1	1	−1
4	43	37	6	4	4
5	58	51	7	5	5
6	56	43	13	6	6
7	76	80	−4	3	−3
8	65	82	3	2	2

மேலுள்ள அட்டவணையிலிருந்து எதிர் தரத்தின் கூட்டல் $1 + 3 = 4$ ஆகும். இது நேர் தரத்தின் கூட்டலைவிட மிகச் சிறியதாகும். இங்கு $N = 8$ ஆகும். அட்டவணை G-லிருந்து $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் $N = 8$-க்கான *T*-ன் தீர்வு கட்டமான மதிப்பு 4 ஆகும். கண்டறிந்த மதிப்பு 4 ஆகும். இது அட்டவணையின் மதிப்புக்குச் சமமாக இருக்கிறது. எனவே H_0-ஐ $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் இருமுனைச் சோதனையின்கீழ் மறுக்கிறோம்.

$N > 25$ எனில் அக் கூறுகளைப் பெரிய கூறுகள் (large samples) என்கிறோம். பெரிய கூறுகளுக்கு அட்டவணை G-ஐப் பயன்படுத்த

முடியாது. இருப்பினும், இம்மாதிரி நிலையில் தரங்களின் கூடுதலான T ஆனது இயல்நிலைப் பரவலாக

$$\text{சராசரி} = \mu_T = \frac{N(N+1)}{4},$$

$$\text{திட்டவிலக்கம்} = \sigma_T = \sqrt{\frac{N(N+1)(2N+1)}{24}}$$

என்று பெற்று அமையும் என்பதை அறியலாம்.

$$\text{எனவே} \quad z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \dfrac{N(N+1)}{4}}{\sqrt{\dfrac{N(N+1)(2N+1)}{24}}}$$

ஆனது தோராயமாக $N(0, 1)$ என்று இயல்நிலைப் பரவலாகி அமையும்.

எனவே, அட்டவணை A-ஐப் பயன்படுத்தி H_0-ன் கீழுள்ள நிகழ்ச்சிக்கியைந்த நிகழ்தகவுகளின் மதிப்புகளைப் பெற்று உய்த்துணரலாம். சிறிய கூறுகளாக இருந்தாலும் இத் தோராயம் மிகச் சிறந்த தோராயம் என்பதை மேலே கொடுக்கப்பட்டுள்ள மாதிரிக் கணக்கைக் கொண்டே அறியலாம். மேலுள்ள மாதிரிக் கணக்கில் கண்டறிந்தவை $T = 4$, $N = 9$ ஆகும்.

$$\text{எனவே} \quad z = \frac{4 - \dfrac{8(9)}{4}}{\sqrt{\dfrac{(8)(9)(17)}{24}}} = -1{\cdot}96$$

அட்டவணை A-விருந்து H_0-ன் கீழுள்ள நிகழ்ச்சிக்கு இயைந்த $z = -1{\cdot}96$-க்கு ஏற்ற நிகழ்தகவு $p = 2(0{\cdot}025) = 0{\cdot}05$ ஆகும். இது மேலுள்ள மாதிரிக் கணக்கில் α என்ற மிகைத்தன்மை மட்டமாகும். இவ்வாறாகச் சிறிய கூறுகளுக்கும் இயல்நிலைத் தோராயத்தைப் பயன்படுத்தலாம் என்பதை அறிகிறோம்.

மாதிரி : 30 ஓட்டுநர்களைக் கொண்ட ஒரு ராண்டம் மாதிரி தேர்ந்தெடுக்கப்பட்டு அவர்கள் மதுபானம் குடிப்பதற்கு முன் பிருந்த பிரதிகிரியை நேரத்தின் (reaction time) அளவும்,

குடித்த பிறகு இருந்த பிரதிகிரியை நேரத்தின் அளவும் கீழ்க் கண்டவாறு அட்டவணைப்படுத்தப்பட்டன.

ஓட்டுநர்கள்	முன்பு	பின்பு
1	62	64
2	65	65
3	61	61
4	63	62
5	66	66
6	64	64
7	67	63
8	72	68
9	70	69
10	69	68
11	72	67
12	70	67
13	75	70
14	68	65
15	64	65
16	69	68
17	70	71
18	73	68
19	63	55
20	66	64
21	69	67
22	68	66
23	64	67
24	65	67
25	70	69
26	70	66
27	68	60
28	72	70
29	71	68
30	70	71

பிரதிகிரியை நேரத்தை மதுபானம் பாதிக்கிறதா என்பதை வில்காக்ஸன் சோதனை கொண்டு காண்போம்.

இப்பொழுது ஓட்டுநர்களின் பிரதிகிரியை நேரத்தை மது பானம் பாதிக்கவில்லை என்ற சூனிய எடுகோளையும், பாதிக்கிறது என்ற மாற்று எடுகோளையும் கொள்வோம். மிகைத்தன்மை மட்டம் $\alpha = 0{\cdot}01$ என்க. இங்குப் பெரிய கூறு அளவையாக உள்ளதால் இயல்நிலைத் தோராயத்தைப் பயன்படுத்த வேண்டியுள்ளது. இப்பொழுது d-க்களைக் கணக்கிட்டு, அதற்கான தரங்களையும் குறிப்பிட்டு, அவைகளில் எந்தக் குறியின் தரங்களின் கூட்டல் குறைவாக உள்ளது என்று கண்டு T-ஐக் காண்போம்.

ஓட்டுநர்கள்	d	d-க்கான தரம்	ஓட்டுநர்கள்	d	குறிகளுடன் d-க்கான தரம்
1	−2	−11·5	16	1	4·5
2	0	—	17	−1	−4·5
3	0	—	18	5	23·0
4	1	4·5	19	8	25·5
5	0	—	20	2	11·5
6	0	—	21	2	11·5
7	4	20·0	22	2	11·5
8	4	20·0	23	−3	−16·5
9	1	4·5	24	−2	−11·5
10	1	4·5	25	1	4·5
11	5	23·0	26	4	20·0
12	3	16·5	27	8	25·5
13	5	23·0	28	2	11·5
14	3	16·5	29	3	16·5
15	−1	−4·5	30	−1	−4·5

மேலுள்ள அட்டவணையின் எதிர் தரங்களின் கூட்டல் நேர் தரங்களின் கூட்டலைவிடக் குறைவாக இருப்பதைக் காணலாம்.

எனவே, எதிர் தரங்களின் கூட்டல் :

$$11{\cdot}5 + 4{\cdot}5 + 4{\cdot}5 + 16{\cdot}5 + 11{\cdot}5 + 4{\cdot}5 = 53{\cdot}0$$

இங்கு 4 ஒத்தவை (ties) உள்ளன.

எனவே $N = 30 - 4 = 26$ ஆகும்.

$$\text{இப்பொழுது, } z = \frac{T - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

$$= \frac{53 - \frac{(26)(27)}{4}}{\sqrt{\frac{(26)(27)(53)}{24}}}$$

$$= -3 \cdot 11$$

$z \cap N(0, 1)$ என்பதால் அட்டவணை A-ல் $\alpha = 0 \cdot 01$ மிகைத் தன்மை மட்டத்தில் z-ன் ஒரு முனையின் மதிப்பான $-3 \cdot 11$-க்கு ஏற்ற $p = 0 \cdot 0009$ ஆகும்.

இந்த $p = 0 \cdot 0009 < \alpha = 0 \cdot 01$ என்பதால் z-ன் மதிப்பானது மறுக்கத்தகு வெளியில் இருப்பதை அறிகிறோம். எனவே, H_0-ஐ மறுக்கிறோம். அதாவது ஓட்டுநர்களின் பிரதிகிரியை நேரத்தை மதுபானம் பாதிக்கிறது என்ற மாற்று எடுகோளை ஏற்றுக்கொள்கிறோம்.

வால்ஷ் சோதனை (Walsh test)

இரு நடத்துகைகளுக்கு, அவைகளின் ஜோடிகளின் மதிப்புகளுக்கிடையேயுள்ள வித்தியாசமான d_i-கள் சமச்சீரான பரவலிலிருந்து வரப்பெற்றவை என்ற தற்கோளை இச் சோதனையில் பயன்படுத்துகிறோம். இதனால் சராசரியானது மையப்போக்கு அளவையின் மிகத் துல்லியமான அளவாக இடைநிலைக்குச் சமமாக இருக்கும். இச் சோதனையில் d_i-கள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்றோ இயல்நிலை முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்றோ தற்கோள்களைக் கொள்ளவில்லை என்பது குறிப்பிடத்தக்கதாகும்.

N ஜோடிகளில் ஒவ்வொரு ஜோடியின் மதிப்புகளுக்கிடையேயுள்ள வித்தியாசமான d_i-களை முதலில் கண்ட பிறகு இவைகளை அவைகளின் அளவுகளுக்கு ஏற்ப வரிசைப்படுத்தி அமைக்கிறோம். d_1 என்பது வித்தியாசங்களில் மிகக்குறைந்த மதிப்பு என்க. இதே போன்று d_2 என்பது வித்தியாசங்களில் அடுத்த குறைந்த மதிப்பு என்க. இவ்வாறாக,

$d_1 \leqslant d_2 \leqslant d_3 \leqslant d_4 \ldots \leqslant d_N$ என்று அமைத்துக் கொள்கிறோம்.

இப்பொழுது இந்த d_i-கள், இடைநிலை மதிப்பைப் பெற்ற முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளைக் கருதுகிறோம். ஒரு சமச்சீர் பரவலில் சராசரியும், இடைநிலையும் ஒன்றி இருக்கும்.

d_i-கள் சமச்சீர் பரவலிலிருந்து வரப்பெற்றவை என்ற தற்கோளை வால்ஷ் சோதனை கருதுகிறது. எனவே μ_0 என்ற வித்தியாச மதிப்புகளின் சராசரி பூச்சியம் என்ற சூனிய எடுகோளையும், இரு முனைச் சோதனைக்கு $\mu_1 \neq 0$ என்ற மாற்று எடுகோளையும், ஒரு முனைச் சோதனைக்கு $\mu_1 > 0$ அல்லது $\mu_1 < 0$ என்ற மாற்று எடுகோளையும் கருதுகிறோம். அட்டவணை H-ஐப் பயன்படுத்தி வால்ஷ் சோதனையின்கீழ் பல்வேறு முடிவுகளுக்கு மிகைத் தன்மையை அறியலாம். இவ்வட்டவணையைப் பயன்படுத்தக் கண்டறிந்த N-ன் மதிப்பும், மாற்று எடுகோளும், d_i-களின் எண்ணுக்குரிய மதிப்பும் (numerical value) தெரியவேண்டும். இவ்வட்டவணை ஒருமுனைச் சோதனைக்கும், இருமுனைச் சோதனைக்குமான மிகைத்தன்மை மதிப்புகளைக் கொடுக்கிறது. இவ்வட்டவணையைப் பயன்படுத்தும் முறையைக் கீழ்க்காணும் உதாரணங்கள் மூலம் காண்போம்.

(i) $N = 5$ என்க; இரு முனைச் சோதனையில் H_0-ஐ $\alpha = 0{\cdot}125$ மிகைத்தன்மை மட்டத்தில் $\frac{1}{2}(d_4 + d_5)$ ஆனது பூச்சியத்தைவிடக் குறைவாகவோ அல்லது $\frac{1}{2}(d_1 + d_2)$ பூச்சியத்தைவிட அதிகமாக இருந்தாலோ H_0-ஐ மறுக்கிறோம்.

மேலும் $\alpha = 0{\cdot}062$ என்ற மிகைத்தன்மை மட்டத்தில் $d_5 < 0$ என்றிருந்தாலோ அல்லது $d_1 > 0$ என்றிருந்தாலோ H_0-ஐ மறுக்கலாம். இப்பொழுது $N = 5$ என இருந்து வித்தியாச மதிப்புகள் பூச்சியத்தைவிட அதிகமாக இருப்பதாக ஊகிப்பதால் $\frac{1}{2}(d_1 + d_2)$ பூச்சியத்தைவிட அதிகமாக இருந்தால், $\alpha = 0{\cdot}062$ மிகைத்தன்மை மட்டத்தில் H_0-ஐ மறுக்கிறோம்.

மேலும் $d_1 > 0$ என்றிருந்தாலும், H_0-ஐ $0{\cdot}031$ மிகைத் தன்மை மட்டத்தில் மறுக்கிறோம். இப்பொழுது நம்முடைய வித்தியாச மதிப்புகள் எதிராக (negative) இருப்பதாக ஊகித்தால், அதாவது $\mu_1 < 0$ என்ற மாற்று எடுகோளைக் கருதினால், $N = 5$ என்றிருந்து, $\frac{1}{2}(d_4 + d_5) < 0$ என்றிருந்தால் H_0-ஐ $\alpha = 0{\cdot}062$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

மேலும் $d_5 < 0$ எனில் H_0-ஐ $\alpha = 0{\cdot}031$ என்ற மிகைத் தன்மை மட்டத்தில் மறுக்கிறோம்.

மாதிரி : 15 மாணவர்கட்கு 10 அர்த்தமற்ற சொற்களின் பாகங்கள் கற்றுக் கொடுக்கப்பட்டன. பின்பு 48 மணி நேரங்கள் கழித்து அவர்கள் கற்றுக்கொண்ட சொற்களைத் திரும்பவும் கூறு மாறு சொன்னதில் அவர்கள் ஞாபகத்தோடு உச்சரித்த அதிர்ச்சி யுடைய சொற்களின் எண்ணிக்கையும் அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கையும் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டன.

மாணவர்கள்	அதிர்ச்சியுடைய சொற்களின் எண்ணிக்கை	அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கை
1	6	3
2	5	3
3	4	0
4	5	3
5	2	3
6	4	2
7	2	3
8	3	2
9	4	1
10	4	3
11	3	4
12	2	3
13	5	2
14	3	4
15	1	0

மேலுள்ள விவரங்களைக் கொண்டு வால்ஷ் சோதனையைப் பயன்படுத்தி நினைவுகூர்ந்து உச்சரிக்கப்பட்ட அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கையானது, அதிர்ச்சியுள்ள சொற்களின் எண்ணிக்கை அதிகமாக மிகைத்தன்மை (signify) செய்கிறதா என்பதைக் காண்போம்.

நினைவுகூர்ந்த அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கைக்கும், அதிர்ச்சியுள்ள சொற்களின் எண்ணிக்கைக்கும் உள்ள இடைநிலை வித்தியாசங்கள் பூச்சியம் என்ற சூனிய எடுகோளையும், நினைவு கூர்ந்த அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கையானது நினைவு கூர்ந்த அதிர்ச்சியுள்ள சொற்களின் எண்ணிக்கையைவிட அதிக மாக இருக்கும் என்ற மாற்று எடுகோளையும் கொள்வோம்.

இப்பொழுது ஒவ்வொரு ஜோடியிலுள்ள மதிப்புகளுக்கிடையே யுள்ள வித்தியாசமான d-க்களை அட்டவணைப் படுத்துவோம்.

மாணவர்கள்	அதிர்ச்சியுடைய சொற்களின் எண்ணிக்கை	அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கை	d
1	6	3	3
2	5	3	2
3	4	0	3
4	5	3	2
5	2	3	−1
6	4	2	2
7	2	3	−1
8	3	2	1
9	4	1	3
10	4	3	1
11	3	4	−1
12	2	3	−1
13	5	2	3
14	3	4	−1
15	1	0	1

மேலுள்ள அட்டவணையில் d-ன் மிகச்சிறிய மதிப்பு -1 ஆகும். இது 5 முறை நிகழ்கிறது. எனவே, $d_1 = -1$, $d_2 = -1$, $d_3 = -1$, $d_4 = -1$, $d_5 = -1$. d-ன் அடுத்த சிறிய மதிப்பு 1 ஆகும். இங்கு இது 3 முறை நிகழ்கிறது. எனவே, $d_6 = 1$, $d_7 = 1$, $d_8 = 1$. இதேபோன்று அடுத்த d-ன் சிறிய மதிப்பான 2 மூன்று முறை நிகழ்கிறது. எனவே, $d_9 = 2$, $d_{10} = 2$, $d_{11} = 2$. அடுத்த சிறிய மதிப்பான 3 ஆனது நான்கு முறை நிகழ்கிறது. எனவே, $d_{12} = 3$, $d_{13} = 3$, $d_{14} = 3$, $d_{15} = 3$ ஆகும். இப்பொழுது அட்டவணை H ஆனது $N = 15$-க்கு $\mu > 0$ என்ற H_1-ன் ஒரு முனைச் சோதனையின்கீழ் $\alpha = 0{\cdot}047$ என்ற மிகைத்தன்மை மட்டத்திற்கு,

$$\text{நீசம்}\ [\tfrac{1}{2}(d_1 + d_{12}), \tfrac{1}{2}(d_2 + d_{11})] > 0$$

என்று காட்டுகிறது.

இங்கு $d_1 = -1$, $d_{12} = 3$, $d_2 = -1$, $d_{11} = 2$ ஆகும்.

எனவே, $\text{நீசம்}\ [\tfrac{1}{2}(-1 + 3), \tfrac{1}{2}(-1 + 2)]$

$= \text{நீசம்}\ [\tfrac{1}{2}(2), \tfrac{1}{2}(1)]$

$= \tfrac{1}{2}(1).$

இங்கு $\frac{1}{2}(1) = \frac{1}{2} > 0$ என்பதால் H_0-ஐ $\alpha = 0 \cdot 047$ மிகைத் தன்மை மட்டத்தில் மறுக்கிறோம்.

அதாவது நினைவு கூர்ந்து உச்சரிக்கப்பட்ட அதிர்ச்சியற்ற சொற்களின் எண்ணிக்கையானது அதிர்ச்சியுள்ள சொற்களின் எண்ணிக்கை அதிகமாக மிகைத்தன்மை செய்கிறது என்று உய்த் துணருகிறோம்.

பயிற்சி 2

1. தேர்தல் பிரசாரம் செய்வதற்கு முன்பு கட்சி A, கட்சி B ஆகியவைகளைத் தேர்ந்தெடுக்கும் நோக்கத்தில், ராண்டம் மாதிரி யிலான 150 நபர்கள் தங்கள் ஓட்டு விருப்பங்களைத் தெரிவித்துக் கொண்டனர். தேர்தல் பிரசாரம் முடிந்த பிறகு அதே நபர்கள் தங்கள் ஓட்டு விருப்பங்களைத் தெரிவித்துக் கொண்டனர். இவ் விவரங்களைக் கீழ்க்காணும் அட்டவணையில் காணலாம்.

		பிரசாரத்திற்குப் பிறகு		மொத்தம்
		கட்சி A	கட்சி B	
பிரசாரத்திற்கு முன்பு	கட்சி A	30	72	102
	கட்சி B	15	33	48

'மெக்னமாரின்' சோதனையைப் பயன்படுத்தி, தேர்தல் பிரசாரத்தினால் ஓட்டு அளிப்பில், ஏதாவது மாறுதல் ஏற்பட் டுள்ளதா என்பதை ஆய்க.

2. X, Y என்ற இரு நிலைகளில் 30 கடினமான வார்த்தை களை மனப்பாடம் செய்வதற்கு எந்தச் சூழ்நிலை சிறந்தது என அறிவதற்கு நடத்தப்பட்ட ஆராய்ச்சியில், 12 பள்ளி மாணவர் களை இவ்விரு சூழ்நிலைகளிலும் ஆராய்ந்ததில் அம் மாணவர்கள் சோதனையில் பெற்ற மதிப்புகள் கீழே கொடுக்கப்பட்டுள்ளன. சூழ்நிலை X, சூழ்நிலை Y-ஐக் காட்டிலும் சிறந்தது என்ற எடுகோளைக் குறியிடு சோதனை மூலம் ஆய்க.

பள்ளி மாணவர்கள்	சூழ்நிலை *X*-ன் மதிப்புகள்	சூழ்நிலை *Y*-ன் மதிப்புகள்
1	16	18
2	50	30
3	32	26
4	37	8
5	11	11
6	12	7
7	19	18
8	16	14
9	41	23
10	17	9
11	13	10
12	14	6

3. 6 மாணவர்கள் தங்களின் எடைகளைக் குறைத்துக்கொள்ளும் நோக்கத்தில் கட்டுப்பாட்டுடன் உணவு உட்கொண்ட பிறகு பெற்ற எடைகளும், உணவுக் கட்டுப்பாட்டிற்கு முன்பு பெற்றிருந்த எடைகளும் பின்வருமாறு:

மாணவர்கள்	முன்பு இருந்த எடை	பின்பு இருந்த எடை
a	174	165
b	191	186
c	188	183
d	182	178
e	201	203
f	188	181

எடையை இழப்பதற்கு உணவுக் கட்டுப்பாடு முக்கிய காரணமாக இருக்கிறதா என்பதைக் குறியிடு சோதனை மூலம் ஆய்க.

4. ஒரே மாதிரியான உறுப்பினர்களுடைய ஒரு ஜோடியில் ஓர் உறுப்பினருக்குச் சிறப்பான பயிற்சி அளிக்கப்படுகிறது. மற்றோர் உறுப்பினருக்கு இப் பயிற்சி அளிக்கப்படவில்லை. இம் மாதிரி 10 ஜோடிகளைக் கொண்ட ஒரு குழுவில் உறுப்பினர்கள் ஒரு சோதனையில் பெற்ற மதிப்புகள் பின்வருமாறு :

ஜோடிகள்	பயிற்சி பெற்றவர்களின் மதிப்புகள்	பயிற்சி பெறாதவர்களின் மதிப்புகள்
1	16	13
2	19	16
3	11	11
4	16	17
5	19	19
6	13	11
7	16	13
8	17	14
9	15	13
10	23	20

பயிற்சி பெற்றவர்கள் பயிற்சி பெறாதவர்களைவிடச் சிறந்தவர்களா என்பதைக் குறியிடு சோதனையைப் பயன்படுத்தி ஆய்க.

5. மாமல்லபுரத்திற்கு 1955ஆம் ஆண்டு முதல் 1973ஆம் ஆண்டுவரை ஒவ்வோர் ஆண்டும் வந்த உல்லாசப் பயணிகளின் எண்ணிக்கை கீழ்க்கண்டவாறு பதிவு செய்யப்பட்டது.

46, 48, 40, 35, 45, 56, 35, 57, 38, 59, 41, 34, 46, 39, 42, 38, 50, 38, 42.

உல்லாசப் பயணிகளின் வருகைக்கான போக்கு வருடம் ஆக ஆக உயர்ந்துகொண்டு செல்கிறதா என்பதை 'காக்ஸ்-ஸ்டூவர்ட்' சோதனை மூலம் ஆய்க.

6. ஒவ்வொரு மாதமும் மேட்டூர் நீர்த்தேக்கத்தில் இருந்து விளைநிலங்களுக்கு ஒரு வினாடிக்கு விடப்படும் தண்ணீரின் வீதம் 24 மாதங்களுக்குக் கன அடியில் பின்வருமாறு குறிக்கப்பட்டது.

மாதம்	முதல் வருடம்	இரண்டாம் வருடம்
ஜனவரி	24·6	25·6
பிப்ரவரி	22·3	20·4
மார்ச்	105·1	125·2
ஏப்ரல்	225·2	195·1
மே	115·4	132·3
ஜூன்	88·4	99·2
ஜூலை	103·2	97·1
ஆகஸ்ட்	85·1	82·2
செப்டம்பர்	86·3	76·2
அக்டோபர்	56·4	49·4
நவம்பர்	32·3	29·6
டிசம்பர்	18·4	16·1

'காக்ஸ்-ஸ்டுவர்டின்' போக்குச் சோதனையைப் பயன்படுத்தி, தண்ணீர் விடப்படும் வீதம் குறையவில்லை என்ற சூனிய எடுகோளுக்கு எதிராகத் தண்ணீர் விடப்படும் வீதம் குறைகிறது என்ற மாற்று எடுகோளைக் கருதி ஆய்க.

7. X, Y என்ற இருவகைப்பட்ட மருந்துகள் 12 நோயாளிகளுக்கு அளிக்கப்பட்டு அதனால் ஏற்பட்ட பிரதிகிரியைகளின் (reactions) மதிப்புகளுக்கு நேர் ஒட்டுறவு உள்ளதா என்பதைக் கீழ்க்கண்ட விவரங்களைக்கொண்டு 'காக்ஸ்-ஸ்டூவர்ட்' போக்குச் சோதனையின் மூலம் அறிக.

நோயாளி	மருந்து X	மருந்து Y
1	0·8	+1·8
2	−1·7	+0·9
3	−0·3	+1·2
4	−1·3	+0·2
5	−0·2	−0·2
6	+3·5	+4·3
7	+3·9	+5·4
8	+0·9	+1·5
9	0	+4·7
10	+2·1	+3·4
11	+1·2	−4·9
12	+1·6	+4·6

8. ஒரு விளைநிலத்தில் 10 இடங்களில் ஒரே மாதிரியான இரு விதைகள் விதைக்கப்பட்டன. முதல் விதையை விதைத்த 12 மணி நேரத்திற்குப் பிறகு இரண்டாம் விதை விதைக்கப்படுகிறது. அவைகளின் விளை பலன்கள் பின்வருமாறு:

இடங்கள்	முதலில் விதைக்கப் பட்ட விதையின் பலன்	இரண்டாவதாக விதைக்கப் பட்ட விதையின் பலன்
1	42	35
2	59	44
3	63	68
4	41	71
5	57	54
6	49	48
7	54	63
8	58	59
9	38	69
10	72	39

இவ்வாறு விளைவிப்பதால், இரண்டாவதாக விதைக்கப்பட்ட விதையைவிட, முதலில் விதைக்கப்பட்ட விதை, அதிகமான பலனைத் தருகிறதா என்பதை 'வில்காக்ஸன்' சோதனையைப் பயன்படுத்தி ஆய்க.

9. ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்ட 20 குமாஸ்தாக்கள் புகை பிடிப்பதற்கு முன்பு பெற்றிருந்த வேலைத் திறன்களும் (மணி, நேரங்களில்), புகை பிடித்த பின்பு பெற்றிருந்த வேலைத் திறன்களும் (மணி, நேரங்களில்) கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டன.

குமாஸ்தாக்கள்	புகை பிடிப்பதற்கு முன்பு	புகை பிடித்த பின்பு
1	5·0	6·5
2	4·5	5·5
3	6·0	5·5
4	4·0	6·5
5	5·0	7·0
6	4·0	5·0
7	5·5	5·5
8	4·5	5·5
9	5·5	6·5
10	6·5	5·5
11	4·5	4·0
12	3·5	4·5
13	4·5	6·0
14	5·5	6·0
15	6·5	5·5
16	6·0	6·5
17	3·5	7·0
18	3·5	7·5
19	4·5	7·5
20	5·5	7·0

புகை பிடிப்பதனல் வேலைத்திறன் பாதிக்கப்படுகிறதா இல்லையா என்பதை 'வில்காக்ஸன்' சோதனையைப் பயன்படுத்தி ஆய்க.

10. 18 மாணவர்களுக்குப் புள்ளியியல் பாடம் கொள்கை (theory) அளவிலும், கணித (mathematical) அளவிலும் கற்றுக் கொடுக்கப்பட்டது. ஒரு மாதம் கழித்து அவர்களுக்கு அப் பாடத்தில் நடத்திய தேர்வில், கொள்கை அளவில் பெற்ற மதிப் பெண்களும், கணித அளவில் பெற்ற மதிப்பெண்களும் பின் வருமாறு :

மாணவர்கள்	கொள்கை அளவில் பெற்ற மதிப்பெண்கள்	கணித அளவில் பெற்ற மதிப்பெண்கள்
1	45	64
2	48	54
3	54	53
4	38	48
5	35	52
6	46	56
7	51	38
8	49	35
9	54	49
10	39	68
11	44	75
12	56	68
13	54	82
14	35	45
15	37	63
16	61	76
17	47	39
18	49	48

மேலுள்ள விவரங்களைக்கொண்டு 'வால்ஷ்' சோதனையைப் பயன்படுத்தி கணித அளவில் பெற்ற மதிப்பெண்கள், கொள்கை அளவில் பெறப்பட்ட மதிப்பெண்களை மிகைத்தன்மை செய்கிறதா என்பதை அறிக.

11. 20 ஓட்டுநர்களுக்கு ஓட்டுநர் பயிற்சிப் பள்ளியில், இலேசான வாகனங்களை ஓட்டவும், கனரக வாகனங்களை ஓட்டவும் பயிற்சி அளிக்கப்பட்டது. ஒரு மாதம் கழித்து அந்த ஓட்டுநர் களுக்கு, வாகனங்களை ஓட்டும் தேர்வு நடத்தப்பட்டதில் அவர்கள் பெற்ற மதிப்பெண்கள் பின்வருமாறு :

ஓட்டுநர்கள்	இலேசான வாகனங்கள்	கனரக வாகனங்கள்
1	35	96
2	92	97
3	45	88
4	86	59
5	39	77
6	42	67
7	68	66
8	94	59
9	98	34
10	56	40
11	49	73
12	57	33
13	63	32
14	72	44
15	68	65
16	41	74
17	39	80
18	36	81
19	43	50
20	51	53

மேலுள்ள விவரங்களைக் கொண்டு 'வால்ஷ்' சோதனையைப் பயன்படுத்தி இலேசான வாகனங்களை ஓட்டிப் பெற்ற மதிப்பெண்கள், கனரக வாகனங்களை ஓட்டிப் பெற்ற மதிப்பெண்களை மிகைத் தன்மை செய்கிறதா என்பதை ஆய்க.

3. இரு சார்பற்ற கூறுகளுக்கான சோதனைகள்

(Tests for Two Independent Samples)

இரு தொடர்புடைய கூறுகளைப் பயன்படுத்துவது நடைமுறையில் முடியாமற்போனாலோ அல்லது பொருத்தமற்றதாகக் காணப்பட்டாலோ இரு சார்பற்ற கூறுகளைப் பயன்படுத்துகிறோம். இவ்விரு சார்பற்ற கூறுகள் ஒவ்வொன்றையும், ராண்டம் முறையில் இரு முழுமைத் தொகுதியிலிருந்து தேர்ந்தெடுக்கப்பட்டதாகவோ அல்லது யாதாமொரு (arbitrary) மூலங்களிலிருந்து (origin) பெறப்பட்ட மாதிரியின் உறுப்புகளுக்கு (members) ராண்டம் முறையில் நிர்ணயிக்கப்பட்ட நடத்துகைகளிலிருந்து வரப்பெற்றதாகவோ இருக்கலாம். இவ்விரு வழிமுறைகளில் எடுக்கப்படும் இரு கூறுகளின் அளவுகள் சமமாக இருக்கவேண்டிய அவசியமில்லை.

உதாரணமாக, கல்லூரியில் படிக்கும் மாணவர்களின் தொகுதியிலிருந்து ஒவ்வோர் உயர் வகுப்பைச் சார்ந்த 7 மாணவர்களையும், ஒவ்வொரு கீழ் வகுப்பைச் சேர்ந்த 11 மாணவர்களையும் தேர்ந்தெடுத்தால் இரு கூறுகளின் அளவுகளும் சமமாக இல்லை.

சுட்டுறுப்பைச் சார்ந்துள்ள முறைகளில் இரு சார்பற்ற கூறுகளைப் பகுப்பாய்வு செய்வதற்கு இரு குழுக்களுக்கான சராசரிக்கு ஸ்டூடன்ஸ் *t* சோதனையைப் பயன்படுத்துவது வழக்கம். *t* சோதனையானது ஒரே மாறுபாடுகளைக் கொண்ட இயல்நிலையைப் பெற்ற முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற தற்கோள்களைக் (assumptions) கொள்கிறோம். ஆனால், சுட்டுறுப்பைச் சாராத முறைகளில் கண்டறிந்த விவரங்களுக்கு இம்மாதிரியான தற்கோள்கள் பொருத்தமுடையதாக இருக்காது. இம்மாதிரியான தருணங்களில் இரு சார்பற்ற கூறுகளுக்கான சுட்டுறுப்பைச் சாராத சோதனைகளைப் பயன்படுத்த வேண்டிய நிலையுள்ளது. அச் சோதனைகளை இவ்வத்தியாயத்தில் காண்போம்.

'ஃபிஷரி'ன் சரியான நிகழ்தகவுச் சோதனை (Fisher's Exact Probability Test)

சார்பற்ற இரு கூறுகளின் அளவுகளும் மிகச் சிறியதாக இருந்து தொடர்ச்சியற்ற (discrete) பெயரளவான அல்லது தர அளவான விவரங்களுக்கு இச் சோதனையைப் பயன்படுத்தலாம்.

இரு சார்பற்ற மாதிரிகளிலிருந்து பெறப்பெற்ற மதிப்புகளுக்கெல்லாம் ஒன்றையொன்று விலக்கும், இரு வகுப்புகளில் ஏதாவதொரு வகுப்பில் இருந்தாலும் இச் சோதனையைப் பயன்படுத்துகிறோம். மதிப்புகளை அலைவெண்களாக 2 நிரைகளையும் (rows) 2 நிரல்களையும் (columns) பெற்ற இணைப்புப் பட்டியலில் குறிக்கிறோம். *A*, *B* என்ற இரு குழுக்கள், உதாரணமாக ஆண்கள், பெண்கள்; தந்தைகள், தாய்மார்கள் என்ற சார்பற்ற குழுக்களாக இருக்கலாம். பொதுவாக இந்த இணைப்பு அட்டவணையின் நிரலின் தலைப்பானது இருவிதமாக —, + என்ற குறிகளாகக் குறிக்கப்படுகிறது. உதாரணமாக + என்பது தேர்ச்சியடைந்த நிலையையும், — என்பது தோல்வி பெற்ற நிலையையும் குறிக்கலாம்.

இரு பாகுபாடுகளில் (two classifications) அமையக்கூடிய விகிதங்களுக்கு, இவ்விரு குழுக்களுக்குள்ள வேறுபாடு உள்ளதா என்பதை இச் சோதனை மூலம் அறியலாம். இப்பொழுது 2×2 இணைப்புப் பட்டியலைக் கீழ்க்கண்டவாறு அமைப்போம்.

2×2 இணைப்புப் பட்டியல்

	—	+	மொத்தம்
குழு 1	a	b	$a+b$
குழு 2	c	d	$c+d$
மொத்தம்	$a+c$	$b+d$	$N=a+b+c+d$

மேலுள்ள விவரங்களிலிருந்து +, — ஆகிய குறிகளின் விகிதங்களுக்குப் பண்பு (attribute) செய்யப்பட்ட முதற்குழுவும், இரண்டாம் குழுவும் வேறுபட்டு மிகைத்தன்மை செய்கிறதா என்பதை அறியலாம்.

(2 × 2) பட்டியலில் ஓரக் கூட்டல்கள் நிலையானது என்றால் அதிகப் பெருக்கப் பரவலின்படி ஒரு குறிப்பிட்ட தலை எண்களின் கட்டுகளைக் கண்டறிவதற்கான சரியான நிகழ்தகவு

$$p = \binom{a+c}{a}\binom{b+d}{b}\Big/\binom{N}{a+b}$$

$$= \frac{\left[\dfrac{(a+c)!}{a!\,c!}\right]\left[\dfrac{(b+d)!}{b!\,d!}\right]}{\dfrac{N!}{(a+b)!\,(c+d)!}}$$

$$= \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{N!\,a!\,b!\,c!\,d!} \qquad (1)$$

அட்டவணை S-ஐப் பயன்படுத்தி இக் காரணிகளின் மதிப்புகளை எளிதில் பெறலாம்.

(1) லுள்ள சூத்திரத்தைப் பயன்படுத்த இப்பொழுது ஓர் உதாரணத்தைக் காண்போம்.

பட்டியலுக்கேற்றவாறு, $a = 10$, $b = 0$, $c = 4$, $d = 5$, எனில் $a+b = 10$, $c+d = 9$, $a+c = 14$, $b+d = 5$, $N = a+b+c+d = 19$ என்க.

அதாவது

	−	+	மொத்தம்
குழு 1	10	0	10
குழு 2	4	5	9
மொத்தம்	14	5	19

இந்த 19 வகைகளும் நான்கு அறைகளில் அமைவதற்கான சரியான நிகழ்தகவு

$$p = \frac{10!\;9!\;14!\;5!}{19!\;10!\;4!\;5!}$$

$$= 0{\cdot}0108$$

எனவே, H_0-ன் கீழ் இம்மாதிரியான அலைவு எண்களின் பரவலுக்கு நிகழ்தகவு $p = 0{\cdot}0108$ எனக் கொள்கிறோம். மேலுள்ள

பட்டியலிலுள்ள அறைகளின் அலைவெண்கள் ஏதும் பூச்சிய மில்லாமலிருந்து பட்டியலின் ஓரக் கூட்டல்களில் எந்த மாற்றமும் இல்லாமலிருந்தால் H_0-ன் கீழ் பரவலிலிருந்து மிகப் பெரிய விலக்கங்கள் ஏற்படும்.

உதாரணமாகக் கீழ்க்காணும் (2×2) பட்டியலைக் கருதுக.

	−	+	மொத்தம்
குழு 1	1	6	7
குழு 2	4	1	5
மொத்தம்	5	7	12

பட்டியல் (1)

இப் பட்டியலில் எந்த அலைவெண்ணும் பூச்சியமாக இல்லை என்பதை அறிக. ஓரக் கூட்டல்களில் ஏதும் மாற்றம் செய்யாமல் மிகப்பெரிய விலக்கங்களோடு நிகழும் அலைவெண்களைப் பெற்ற கீழ்க்காணும் பட்டியலைக் கருதுவோம்.

	−	+	மொத்தம்
குழு 1	0	7	7
குழு 2	5	0	5
மொத்தம்	5	7	12

பட்டியல் (2)

எனவே, பட்டியல் (1)லுள்ள விவரங்களுக்குச் சூனிய எடுகோளை ஒரு புள்ளியியல் சோதனையைச் செய்ய சூத்திரம் (1) ஐப் பயன்படுத்தி, பட்டியல் (1)-க்கும் பட்டியல் (2)-க்கும் p-க்களைக் கண்டு இவ்விரு p-க்களையும் கூட்டவேண்டும்.

இவ்வாறாக, $p = \dfrac{7!\ 5!\ 5!\ 7!}{12!\ 1!\ 6!\ 4!\ 1!}$

$= 0{\cdot}04399$

$$\text{மேலும் } p = \frac{7!\ 5!\ 5!\ 7!}{12!\ 0!\ 7!\ 0!\ 5!}$$

$$= 0{\cdot}00126$$

$$\text{நிகழ்தகவு } p = 0{\cdot}04399 + 0{\cdot}00126$$

$$= 0{\cdot}04525$$

அதாவது $p = 0{\cdot}04525$ ஆனது பட்டியல் (1) லுள்ள விவரங்கள் H_0-ஐ மறுக்கலாமா என்பதைத் தீர்மானிப்பதற்கு உதவும் மதிப்பாகும். இணைப்புப் பட்டியலிலுள்ள ஏதாவதோர் அலைவெண் மிகச்சிறியதாக இருந்தால் ஃபிஷர் சோதனைக்கான கணக்கீடுகளைக் காண்பது மிகவும் கடினமாகிறது. உதாரணமாக ஓர் அறையின் மிகச்சிறிய அலைவெண்ணின் மதிப்பு 2 எனில், பிறகு மூன்று சரியான நிகழ்தகவு மதிப்புகளை, சூத்திரம் (1)ஐப் பயன்படுத்திக் கண்டபிறகு, முதலில் கண்ட மதிப்பு, 2 உடன் கூட்டவேண்டும். ஓர் அறையின் மிகச்சிறிய அலைவெண்ணின் மதிப்பு 3 என்றால், பிறகு நான்கு சரியான நிகழ்தகவு மதிப்புகளைச் சூத்திரம் (1)ஐப் பயன்படுத்திக் கண்டுபிடித்து முன்னதோடு கூட்டவேண்டும்.

p-ன் சரியான மதிப்பைக் கொள்ளாமல் மிகைத்தன்மை மட்டத்தைக் கொண்டால் போதுமானது என்று கருதினல் அட்டவணை I-ஐப் பயன்படுத்தலாம்.

$N = 30$ அல்லது அதற்குக் குறைவாக உள்ள மதிப்புகளுக்கு மட்டுமே அட்டவணை I-ஐப் பயன்படுத்தமுடியும். மேலும் வலது ஓரக் கூட்டல்களான, $(a + b)$, $(c + d)$ ஆகிய இரு மதிப்புகளில் ஏதேனும் ஒன்று 15-க்குக் குறைவாகவும், மற்றொன்று 15-ஆகவோ அல்லது அதிகமாகவோ இருத்தல் அவசியம். அப்பொழுதுதான் அட்டவணை I-ஐப் பயன்படுத்தமுடியும். சில சமயங்களில் (2×2) பட்டியலில் நிரலின் கூட்டல்களான $(a + c)$, $(b + d)$ ஆகியவை மேற்சொன்ன நிபந்தனைகளைப் பூர்த்திசெய்யும். ஆனால், வலது ஓரக் கூட்டலான (நிரையின் கூட்டல்) மேற்சொன்ன நிபந்தனைகளைப் பூர்த்தி செய்யாத நிலை எழலாம். அம்மாதிரியான நிலைகளில், நிரைகளை, நிரல்களாக மாற்றியமைத்து இணைப்புப் பட்டியலை அமைத்து, அட்டவணை I-ஐப் பயன்படுத்தலாம்.

அட்டவணை I-ஐப் பயன்படுத்தக் கீழ்க்காணும் வழிமுறைகளைப் பின்பற்ற வேண்டும்:

1. முதலில் $a + b$, $c + d$ ஆகிய மதிப்புகளைக் காண வேண்டும்.

2. $(a+b)$ஐ வலது ஓரக் கூட்டல்கள் என்ற தலைப்பில் கண்டு $(c+d)$க் கேற்ற மதிப்பை அட்டவணை I-ல் காண வேண்டும்.
3. $(c+d)$-க்குப் பல மதிப்புகள் கொடுக்கப்பட்டுள்ளன. எனவே $(a+b)$, $(c+d)$ஐக் கருதும்போது b-யையும் கருத்தில்கொண்டு அட்டவணை I-லுள்ள மதிப்பைக் காண முற்படவேண்டும்.
4. இப்பொழுது (2×2) பட்டியலிலுள்ள d-ன் மதிப்பு அட்டவணை I-ல் α என்ற மிகைத்தன்மை மட்ட மதிப்பிற்குக் குறைவாகவோ அல்லது சமமாகவோ இருந்தால் கண்டறிந்த d-ன் மதிப்பு, α-என்ற மிகைத்தன்மை மட்டத்தில் மிகைத்தன்மை செய்கிறது என்கிறோம்.

உதாரணமாக,

$a=1$	$b=8$
$c=6$	$d=0$

என்ற (2×2) பட்டியலைக் கருதுக. அதனால் $c+d=6$, $b=8$, $d=0$ என்ற கண்டறிந்த மதிப்புகளுக்கு அட்டவணை I-ல் $a+b=9$, $c+d=6$, $b=8$-க்கு $\alpha=0{\cdot}05$ என்ற மிகைத் தன்மை மட்டத்தில் அட்டவணையின் மதிப்பு 2 ஆகும். ஆனால் (2×2) பட்டியலில் கண்டறிந்த d-ன் மதிப்பு 0 ஆகும். இது அட்டவணையின் மதிப்பான 2-ஐ விடக் குறைவாக உள்ளதால் கண்டறிந்த விவரங்கள் $\alpha=0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் மிகைத்தன்மை செய்கிறது என்று கொள்கிறோம்.

குறிப்பு : கண்டறிந்த b-ன் மதிப்பை அட்டவணையில் காண முடியாதுபோனல் கண்டறிந்த a-ன் மதிப்பைப் பயன்படுத்தி அட்டவணையில் α-க் கேற்ற மதிப்பைக் காண்கிறோம். பட்டியலிலுள்ள கண்டறிந்த d-ன் மதிப்பிற்குப் பதிலாக c-ன் மதிப்பைக் கண்டு, $c \leqslant$ அட்டவணை மதிப்பு என்றால், α என்ற மிகைத்தன்மை மட்டத்தில் கண்டறிந்த விவரங்கள் மிகைத்தன்மை செய்கிறது என்று கூறுகிறோம்.

மற்றோர் உதாரணமாக (2×2) இணைப்புப் பட்டியல் (1)ஐக் கருதுக. இப் பட்டியலில் $a+b=7$, $c+d=5$, $b=6$, $d=1$ ஆகும். அட்டவணைI-லிருந்து $\alpha=0{\cdot}05$ மிகைத்தன்மை மட்டத்

திற்கான மதிப்பு 1 ஆகும். இது அட்டவணையின் மதிப்பிற்குச் சமமாக உள்ளது. எனவே, $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் ஒரு முனைச் சோதனைக்குக் கண்டறிந்த விவரங்கள் மிகைத்தன்மை செய்கிறது என்று உய்த்துணர்கிறோம். ஆனால் ஃபிஷரின் சரியான நிகழ்தகவிற்கான சூத்திரத்தைப் பயன்படுத்திக் கிடைத்த p-ன் மதிப்பு $0{\cdot}045$ ஆகும். இது $\alpha = 0{\cdot}05$-க்கு நெருங்கியுள்ளது. எனவே, அட்டவணை I-ன் மதிப்புத் தோராயமாக உள்ளது என்பது தெளிவாகிறது.

இருமுனைச் சோதனைக்குக் கண்டறிந்த மிகைத்தன்மை மட்டத்தை இருமடங்காகக் கருதுகிறோம்.

இவ்வாறாக $\alpha = 2\ (0{\cdot}05) = 0{\cdot}10$ மட்டத்தில் பட்டியல் (1)லுள்ள விவரங்கள் H_0-ஐ மறுப்பதற்கு ஏற்றவாறுள்ளன என்று முடிவு செய்கிறோம்.

மாதிரி : ஆண்கள், பெண்கள் ஆகிய இரு குழுவினர் செய்யும் இருவிதமான தொழிலுக்கான அலைவெண்கள் 2×2 இணைப்புப் பட்டியலாகக் கீழ்க்கண்டவாறு கொடுக்கப்பட்டுள்ளன.

	மருத்துவத் தொழில் செய்வோர்	மற்றத் தொழில் செய்வோர்	மொத்தம்
ஆண்கள்	2	8	10
பெண்கள்	5	0	5
மொத்தம்	7	8	15

மேலுள்ள விவரங்களைக் கொண்டு, ஆண்களும், பெண்களும் சமமான விகிதத்தில் தொழில்களை நாடுகின்றனர் என்ற சூனிய எடுகோளையும், பெரும்பாலான ஆண்கள் பெண்களைவிட மற்றத் தொழில்களை நாடுகின்றனர் என்ற மாற்று எடுகோளையும், 'ஃபிஷரின்' சரியான நிகழ்தகவுச் சோதனை மூலம் சோதனை செய்வோம். $N = (a + b + c + d) = 2 + 8 + 5 + 0 = 15$ என்பது சிறிய எண் என்பதால், ஃபிஷரின் சோதனையைச் செய்ய விரும்புகிறோம்.

$\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கொள்வோம். அட்டவணை I-ல் $a + b = 10, c + d = 5, b = 8$-க்கு $\alpha = 0{\cdot}05$-க்கேற்ற மதிப்பு 1 ஆகும்.

கொடுக்கப்பட்டுள்ள பட்டியலிலுள்ள கண்டறிந்த d-ன் மதிப்பு 0 ஆகும். இது அட்டவணை I-ன் மதிப்பான 1ஐ விடக்

குறைவாக உள்ளதால், கண்டறிந்த விவரங்களை மிகைத்தன்மை செய்கிறது என்று உய்த்துணர்கிறோம். அதாவது H_0-ஐ $\alpha = 0 \cdot 05$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். எனவே, ஆண்களும், பெண்களும் தங்கள் தொழில்களை நாடுவதில் வேற்றுமை கொண்டுள்ளனர் என்று உய்த்துணருகிறோம்.

'டாச்சரின்' மாற்று அமைப்பு (Tocher's modification)

'ஃபிஷரின்' சரியான நிகழ்தகவுச் சோதனைக்கு டாச்சரின் மாற்று அமைப்பானது, (2×2) இணைப்புப் பட்டியலுக்கு மிகவும் சக்திவாய்ந்த ஒரு கூறுச் சோதனையை அளிக்கிறது. இந்த மாற்று அமைப்பை ஓர் உதாரணம் மூலம் காண்போம்.

இப்பொழுது கீழுள்ள (2×2) இணைப்புப் பட்டியலைக் கருதுவோம்.

x

2	6	8
4	2	6
6	8	14

பட்டியல் (3)

இப் பட்டியலிலுள்ள குறைந்த அலைவெண்ணுக்கு மிகப் பெரிய வெளிப்பாடுகளைக் கொண்டு (more extreme outcomes) வலது ஓரக்கூட்டலில் மாறுதல் ஏதுமின்றிக் கீழ்க்காணும் பட்டியல்கள் (4), (5) மூலம் குறிப்போம்.

y

1	7	8
5	1	6
6	8	14

பட்டியல் (4)

z

0	8	8
6	0	6
6	8	14

பட்டியல் (5)

கண்டறிந்த விவரங்களைக் கொண்ட பட்டியல் (3)-ஐக் கொண்டு H_0-ஐ $\alpha = 0\cdot05$ மிகைத்தன்மை மட்டத்தில் சோதனை செய்ய விரும்புகிறோம். இப்பொழுது சூத்திரம் (1)-ஐப் பயன் படுத்தி ஒவ்வொரு பட்டியலுக்குமான,

$$p_x = \frac{8!\ 6!\ 6!\ 8!}{14!\ 2!\ 6!\ 4!\ 2!}$$

$$= 0\cdot1398$$

$$p_y = \frac{8!\ 6!\ 6!\ 8!}{14!\ 1!\ 7!\ 5!\ 1!}$$

$$= 0\cdot01598$$

$$p_z = \frac{8!\ 6!\ 6!\ 8!}{14!\ 0!\ 8!\ 6!\ 0!}$$

$$= 0\cdot00033$$

H_0-ன் கீழ் மிகப்பெரிய மதிப்புகள், கண்டறிந்த மதிப்பு களானது ஏற்படுவதற்கு இணங்கிய நிகழ்தகவு $p_x+p_y+p_z$ ஆகும்.

அதாவது $p = p_x + p_y + p_z$

$$= 0\cdot1398 + 0\cdot01598 + 0\cdot00033$$

$$= 0\cdot2000.$$

இவ்வாறாக ஃபிஷரின் சோதனைப்படி $p = 0\cdot2000$ ஆகும்.

'டாச்சரின்' மாற்று அமைப்பில் கண்டறிந்த விவரத்தைத் தவிர வலது ஓரக் கூட்டல்களில் மாற்றம் எதுவுமில்லாதவாறு அமைக்கப்பட்ட மிகப் பெரிய வெளிப்பாடுகளைக் கொண்ட மற்றப் பட்டியல்களுக்கான நிகழ்தகவு மதிப்புகளான p_y, p_z ஆகியவை களைக் கண்டு அவைகளைக் கூட்டிக் கிடைக்கும் நிகழ்தகவான $p_y + p_z$ மிகைத்தன்மை மட்டமான α-ஐ விட அதிகமாக இருந் தால் H_0-ஐ மறுக்கமுடியாது. ஆனால், இந்த நிகழ்தகவு, மிகைத் தன்மை மட்டமான α-ஐ விடக் குறைவாக இருந்து, ஃபிஷர் சோதனைக்கான நிகழ்தகவான $p = p_x+p_y+p_z$ ஆனது α-ஐ விட அதிகமாக இருந்தால்,

$$\frac{\alpha - p_{\text{மிகப் பெரிய வகைகள்}}}{p_{\text{கண்டறிந்த வகை}}}$$

என்ற விகிதத்தைக் காணவேண்டும்.

பட்டியல் (3)-க்கு ஏற்ப,

$$\frac{\alpha-(p_y+p_z)}{p_x}=\frac{0\cdot05-(0\cdot01598+0\cdot00033)}{0\cdot1398}$$

$$=\frac{0\cdot05-0\cdot01631}{0\cdot1398}$$

$$=\frac{0\cdot03369}{0\cdot1398}$$

$\simeq 0\cdot241$ ஆகும்.

இப்பொழுது, ராண்டம் எண்களிலிருந்து 0,1 ஆகிய எண்களுக்கிடையே அமைந்துள்ள ஓர் எண்ணை, ராண்டம் முறையில் தேர்ந்தெடுக்கவேண்டும். இம்மாதிரி, தேர்ந்தெடுக்கப்படும் ராண்டம் எண்ணுனது மேற்கண்ட விகிதத்தைவிடக் குறைவாக இருந்தால், (அதாவது $0\cdot241$-ஐ விடக் குறைவாக இருந்தால்) H_0-ஐ மறுக்கிறோம். அதிகமாக இருந்தால் H_0-ஐ மறுக்க முடியாது.

இரு சார்பற்ற கூறுகளுக்குக் கைவர்க்கச் சோதனை (Chi-square Test for Two Independent Samples)

கொடுக்கப்பட்ட விவரங்களிலுள்ள அலைவெண்கள் தொடர்ச்சியற்ற வகையைச் சார்ந்தனவாக இருந்தால் இரு சார்பற்ற குழுக்களுக்கிடையேயுள்ள வேற்றுமைக்கான மிகைத் தன்மையை கைவர்க்கச் சோதனையைக்கொண்டு அறியலாம். இச் சோதனையில் இரு குழுக்களும் சில சிறப்பியல்புகளுக்கு வேறுபடுகின்றன என்ற சூனிய எடுகோளைக் கொள்கிறோம். இவ்வெடுகோளைச் சோதனை செய்யப் பல்வேறு பிரிவுகளில் (categories) அமையக்கூடிய ஒவ்வொரு குழுவிலுள்ள மொத்த வகைகளைக் (cases) கண்டபின்பு ஒரு குழுவிலுள்ள விகிதப்படியான (proportionated) வகைகளைப் பெற்ற பிரிவுகளை மற்றொரு குழுவிலுள்ள விகிதப்படியான வகைகளுக்கு ஒப்பிட்டுப் பார்க்கிறோம். உதாரணமாக, பல வகைப்பட்ட இசைகளை விரும்புவதில் இரு குழுக்களுக்கிடையே வேற்றுமை உள்ளதா என்று அறியலாம். சூனிய எடுகோளைச் சோதனை செய்ய,

$$\chi^2=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} \qquad (1)$$

என்ற சூத்திரத்தைக் கருதுகிறோம்.

இதில் O_{ij} என்பது கொடுக்கப்பட்ட விவரங்களில் முறையே iஆம் நிரைக்கேற்ற, jஆம் நிரலினுள் கண்டறிந்த அலைவெண்ணாகும்.

E_{ij} என்பது H_0-ன் கீழ் iஆம் நிரைக்கேற்ற, jஆம் நிரலிலுள்ள எதிர்பார்க்கும் அலைவெண்ணாகும்.

(1)லுள்ள கைவர்க்கத்தின் மதிப்புகள் கைவர்க்கப் பரவலாக $(m-1)$ $(n-1)$ பாகைகளைப் பெற்றதாக அமைகிறது. இதில் முறையே m, n என்பன இணைப்புப் பட்டியலிலுள்ள நிரைகளின் எண்ணிக்கையையும், நிரல்களின் எண்ணிக்கையையும் குறிக்கின்றன.

இப்பொழுது கீழ்க்காணும் இணைப்புப் பட்டியலைக் கருதுவோம்.

	குழு 1	குழு 2	மொத்தம்
வகை 1	a_{11}	a_{12}	$a_{11}+a_{12}$
வகை 2	a_{21}	a_{22}	$a_{21}+a_{22}$
வகை 3	a_{31}	a_{32}	$a_{31}+a_{32}$
மொத்தம்	$a_{11}+a_{21}+a_{31}$	$a_{12}+a_{22}+a_{32}$	$a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32}$

மேலுள்ள அட்டவணையில் 3 நிரைகளும், 2 நிரல்களும் உள்ளன. எனவே, இங்கு $m = 3$, $n = 2$ ஆகும். மேலும், இவ்வட்டவணையிலுள்ள அலைவெண்களான a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32} ஆகியவை கண்டறிந்த அலைவெண்களாகும். இப்பொழுது எதிர்பார்க்கும் அலைவெண்களைக் காணக் கீழ்க்காணும் முறையைப் பின்பற்றுகிறோம். வகை (i) குழு (j)க்கான கண்டறிந்த அலைவெண்ணான a_{ij}-க்கு ஏற்ற எதிர்பார்க்கும் அலைவெண்ணை, அது அமைந்துள்ள நிரையின் கூட்டலையும், நிரலின் கூட்டலையும் பெருக்கி அவ்வட்டவணையிலுள்ள மொத்த அலைவெண்ணால் வகுக்க வேண்டும். அவ்வாறே வகை 1, குழு 1-க்கான எதிர்பார்க்கும்

அலைவெண் $$A_{11} = \frac{(a_{11}+a_{21}+a_{31})\,(a_{11}+a_{12})}{(a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32})}$$ ஆகும்.

இதேபோன்று, $A_{12}=\dfrac{(a_{12}+a_{22}+a_{32})\,(a_{11}+a_{12})}{(a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32})}$

$$A_{21}=\frac{(a_{11}+a_{21}+a_{31})\,(a_{21}+a_{22})}{(a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32})}$$

$$A_{22}=\frac{(a_{12}+a_{22}+a_{32})\,(a_{21}+a_{22})}{(a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32})}$$

$$A_{31}=\frac{(a_{11}+a_{21}+a_{31})\,(a_{31}+a_{32})}{(a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32})}$$

$$A_{32}=\frac{(a_{12}+a_{22}+a_{32})\,(a_{31}+a_{32})}{(a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32})}$$

இவ்வாறாக எதிர்பார்க்கும் அலைவெண்களைக் கொண்ட கீழ்க் காணும் இணைப்புப் பட்டியலை அமைக்கிறோம்.

	குழு 1	குழு 2
வகை 1	A_{11}	A_{12}
வகை 2	A_{21}	A_{22}
வகை 3	A_{31}	A_{32}

இப்பொழுது சூத்திரம் (6-9)ஐப் பயன்படுத்தினால்,

$$\chi^2=\sum_{i=1}^{3}\sum_{j=1}^{2}\left(\frac{O_{ij}-E_{ij}}{E_{ij}}\right)^2$$

$$=\frac{(a_{11}-A_{11})^2}{A_{11}}+\frac{(a_{12}-A_{12})^2}{A_{12}}+\frac{(a_{21}-A_{21})^2}{A_{21}}$$

$$+\frac{(a_{22}-A_{22})^2}{A_{22}}+\frac{(a_{31}-A_{31})^2}{A_{31}}+\frac{(a_{32}-A_{32})^2}{A_{32}}$$

$= A$ என்க.

இந்த χ^2 ஆனது $(3-1)\ (2-1) = 2$ வரையற்ற பாகைகளைப் பெற்று கைவர்க்கப் பரவலாக அமைகிறது. இப்பொழுது அட்டவணை C-யைப் பயன்படுத்தி α என்ற மிகைத்தன்மை மட்டத்திற் கேற்ற கைவர்க்க மதிப்புகளைப் பெறலாம். கணக்கிடப்பட்ட χ^2 மதிப்பான A ஆனது அட்டவணை C-லுள்ள, α என்ற மிகைத் தன்மை மட்டத்திற்கேற்ற மதிப்பைவிட அதிகமாக இருந்தால்

α மிகைத்தன்மை மட்டத்தில் வேற்றுமை இல்லை என்ற சூனிய எடுகோளை மறுக்கிறோம்.

மாதிரி: ஒரு வினாவிற்கு விடையளிக்கும்போது மூன்று வகையான மனப்பான்மையைப் பெற்ற மாணவ, மாணவியரின் எண்ணிக்கையும் கீழுள்ள இணைப்புப் பட்டியலில் கொடுக்கப் பட்டுள்ளது.

	ஏற்றுக் கொள்ளுதல்	கவலையின்றி இருத்தல்	மறுத்தல்	மொத்தம்
மாணவர்கள்	12	22	9	43
மாணவிகள்	32	14	6	52
மொத்தம்	44	36	15	95

வினாவிற்கு விடைபற்றிய மனப்பான்மையில் வேறுபாடு மிகைத்தன்மை செய்கிறதா என்பதை ஆய்க.

வினாவிற்கு விடையளிப்பதில் இன வேறுபாடு இல்லை என்ற சூனிய எடுகோளையும், வேறுபாடு இருக்கிறது என்ற மாற்று எடுகோளையும் கொள்வோம். $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். கொடுக்கப்பட்டுள்ள இணைப்புப் பட்டியலிலுள்ள அலைவெண்கள் கண்டறிந்த மதிப்புகளாகும். இப்பொழுது எதிர்பார்க்கும் அலைவெண்களைக் காண்போம்.

$$A_{11} = \frac{44 \times 43}{95} = 19{\cdot}9$$

$$A_{12} = \frac{36 \times 43}{95} = 16{\cdot}3$$

$$A_{13} = \frac{15 \times 43}{95} = 6{\cdot}8$$

$$A_{21} = \frac{44 \times 52}{95} = 24{\cdot}1$$

$$A_{22} = \frac{36 \times 52}{95} = 19{\cdot}7$$

$$A_{23} = \frac{15 \times 52}{95} = 8{\cdot}2$$

இப்பொழுது இந்த எதிர்பார்க்கும் அலைவெண்களை இணைப்புப் பட்டியலாகக் கீழ்க்கண்டவாறு அமைப்போம்.

	ஏற்றுக் கொள்ளுதல்	கவலையின்றி இருத்தல்	மறுத்தல்	மொத்தம்
மாணவர்கள்	19·9	16·3	6·8	43
மாணவிகள்	24·1	19·7	8·2	52
மொத்தம்	44	36	15	95

இப்பொழுது $\chi^2 = \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$

$$= \frac{(12-19\cdot9)^2}{19\cdot9} + \frac{(22-16\cdot3)^2}{16\cdot3} + \frac{(9-6\cdot8)^2}{6\cdot8}$$

$$+ \frac{(32-24\cdot1)^2}{24\cdot1} + \frac{(14-19\cdot7)^2}{19\cdot7} + \frac{(6-8\cdot2)^2}{8\cdot2}$$

$$= 3\cdot14 + 1\cdot99 + 0\cdot71 + 2\cdot59 + 1\cdot65 + 0\cdot59$$

$$= 10\cdot67.$$

இந்த χ^2 ஆனது $(2-1)\ (3-1) = 2$ வரையற்ற பாகைகளைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. அட்டவணை C-லிருந்து $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்திற்கான மதிப்பு 5·99 ஆகும்.

கணக்கிடப்பட்ட மதிப்பான 10·67 ஆனது அட்டவணையின் மதிப்பான 5·99-ஐ விட அதிகமாக உள்ளதால் $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்தில் வினவிற்கு விடை அளிப்பதில் இன வேறுபாடு இல்லை என்ற சூனிய எடுகோளை மறுக்கிறோம்.

	−	+
குழு I	a	b
குழு II	c	d

என்ற 2×2 இணைப்புப் பட்டியலுக்கு இச் சோதனையைப் பயன்படுத்தும்போது, χ^2-ன் மதிப்பைக் காணப் பின்வரும் சூத்திரத்தைப் பயன்படுத்த வேண்டும்.

$$\chi^2 = \frac{N\left(\mid ad - bc \mid - \frac{N}{2}\right)^2}{(a+b)\,(c+d)\,(a+c)\,(b+d)}$$

இச் சூத்திரத்தைப் பயன்படுத்துவது மிகவும் எளிது என்பது மட்டுமல்லாமல், தொடர்ச்சிக்கான திருத்தத்தையும் மேற் கொள்கிறது.

மாதிரி: உற்பத்தி செய்யப்பட்ட பொருள்களைக் கொண்ட இரண்டு வண்டிகளில் பழுதான பொருள்களின் விகிதம் இரு வண்டிச் சுமைகளுக்கிடையே வேறுபடுகிறதா என்பதை அறியும் முகமாக, ராண்டம் முறையில் மாதிரி செய்ததில் முதல் வண்டி யின் சுமையில் 21 பொருள்களில் 10 பழுதற்றனவாகவும், இரண்டாம் வண்டியின் சுமையில் 59 பொருள்களில் 46 பொருள்கள் பழுதற்றவையாகவும் இருந்தன.

	பழுதற்றவை	பழுதுடையவை	மொத்தம்
முதல் வண்டியின் கனம்	10	11	21
இரண்டாம் வண்டியின் கனம்	46	13	59
மொத்தம்	56	24	80

இப்பொழுது இரு வண்டிகளின் சுமைகளுக்கிடையே வேறு பாடு இல்லை என்ற சூனிய எடுகோளையும் பழுதற்ற பொருள் களின் விகிதம், பழுதுடைய பொருள்களின் விகிதத்தை விட அதிக மாக உள்ளது என்ற மாற்று எடுகோளையும் கொண்டு இரு சார் பற்ற கூறுகளுக்கான கைவர்க்கச் சோதனையைச் செய்வோம்.

$$\chi^2 = \frac{N\left(\mid ad - bc \mid - \frac{N}{2}\right)^2}{(a+b)\,(c+d)\,(a+c)\,(b+d)}$$

இதில் $a = 10, b = 11, c = 46, d = 13$

$N = a + b + c + d = 80$

$$\text{எனவே } \chi^2 = \frac{80\left[\mid (10)\,(13) - (11)\,(46) \mid - \frac{80}{2}\right]^2}{(21)\,(59)\,(56)\,(24)}$$

$$= \frac{80\,(336)^2}{1665216}$$

$$= 5{\cdot}42$$

இந்த χ^2 ஆனது $(2-1)$ $(2-1) = 1$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாகும். அட்டவணை C-லிருந்து $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்திற்கு 1 வரையற்ற பாகைக்கான மதிப்பு 3·84 ஆகும்.

5·42 ஆனது அட்டவணையின் மதிப்பான 3·84-யைவிட அதிகமாக உள்ளதென்பதால் $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் சூனிய எடுகோளான H_0-ஐ மறுக்கிறோம். அதாவது இரு வண்டிகளின் சுமைகளுக்கிடையே வேறுபாடு இல்லை என்ற சூனிய எடுகோளை மறுக்கிறோம்.

குறிப்பு: இணைப்புப் பட்டியலில் எதிர்பார்க்கும் அலைவெண்களின் மதிப்புப் பெரியதாக இருந்தால் கைவர்க்கச் சோதனையைச் செய்கிறோம். கண்டறிந்த, எதிர்பார்க்கும் அலைவெண்கள் இந்த நிபந்தனையைப் பூர்த்தி செய்யவில்லை என்றால், அந்த அறைகளிலுள்ள மதிப்புகளை மற்றோர் அறைகளிலுள்ள மதிப்புகளுடன் இணைக்கிறோம். இவ்வாறு செய்யும்போது பட்டியலிலுள்ள விவரங்களில் மாறுதல் எதுவும் ஏற்படா வண்ணம் பார்த்துக் கொள்ளவேண்டும்.

2×2 இணைப்புப் பட்டியலுக்கு $N > 40$ எனில் தொடர்ச்சிக்கான திருத்தத்தைக் கொண்ட கைவர்க்கச் சூத்திரத்தைப் பயன்படுத்துகிறோம். N-ன் மதிப்பு 20-க்கும், 40-க்கும் இடையே இருந்து எதிர்பார்க்கும் அலைவெண்களின் மதிப்பு 5 அல்லது அதற்கு மேற்பட்டிருந்தால் தொடர்ச்சிக்கான திருத்தத்தைப் பெற்ற கைவர்க்கச் சூத்திரத்தைப் பயன்படுத்துகிறோம். எதிர்பார்க்கும் அலைவெண்கள் 5-ஐ விடக் குறைவாக இருந்தாலோ, $N < 20$ என்றாலோ ஃபிஷரின் சோதனையைப் பயன்படுத்த வேண்டும். (2×2) இணைப்புப் பட்டியலைத் தவிர மற்ற இணைப்புப் பட்டியலிலுள்ள அறைகளின் 20 சதவீத, 2 அறைகளிலுள்ள எதிர்பார்க்கும் அலைவெண்கள் 5-க்குக் குறைவாகவும் ஆனால் 1-க்கு அதிகமாகவும் இருந்தால், அந்த நிரையிலுள்ள அல்லது நிரலிலுள்ள அலைவெண்களை நிபந்தனைகள் பூர்த்தி செய்யுமாறு அடுத்த நிரையிலுள்ள அல்லது நிரலிலுள்ள முறையான அறையின் அலைவெண்களோடு கூட்டிய பின்பு பெறப்பட்ட இணைப்புப் பட்டியலுக்கு கைவர்க்கச் சோதனையைச் செய்யவேண்டும்.

இடைநிலைச் சோதனை (Median Test)

மையப் போக்குகளில் இரு குழுக்களுக்கிடையே வேறுபாடு உள்ளதா இல்லையா என்பதை அறிவதே இடைநிலைச் சோதனை

யாகும். அதாவது ஒரே அளவுகளைப் பெற்ற அல்லது பெறாத இரு சார்பற்ற குழுக்கள் ஒரே இடைநிலையைப் பெற்ற முழுமைத் தொகுதியிலிருந்து வரப்பெற்றனவா என்பதைச் சோதனை செய்து அறியலாம். இரு சார்பற்ற குழுக்களும் ஒரே இடைநிலையைப் பெற்ற முழுமைத் தொகுதிகளிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளையும், ஒரு முழுமைத்தொகுதியின் இடைநிலையிலிருந்து வேறுபட்டுள்ளன என்ற இருமுனைச் சோதனைக்கான மாற்று எடுகோளையும் அல்லது ஒரு முழுமைத் தொகுதியின் இடைநிலையின் மதிப்பைவிட அதிகமானவை என்ற ஒரு முனைச் சோதனைக்கான மாற்று எடுகோளையும் அமைத்துச் சோதனை செய்கிறோம். இடைநிலைச் சோதனையைச் செயல்படுத்த, முதலில் இரு குழுக்களையும் இணைத்துக் கிடைத்த தொடருக்கு இடைநிலையைக் காண்கிறோம். பின்பு முதல் குழுவிலும், இரண்டாம் குழுவிலும் தனித் தனியாக எத்தனைக் கண்டறிந்தவை இந்த இடைநிலையின் மதிப்பை விட அதிகமாக உள்ளன, குறைவாக உள்ளன என்று காண்கிறோம். பின்பு கீழ்க்கண்ட (2×2) இணைப்புப் பட்டியலை அமைக்கிறோம்.

	குழு I	குழு II	மொத்தம்
இடைநிலை மதிப்பைவிட அதிகமாக உள்ள மதிப்புகளின் எண்ணிக்கை	p	q	$p+q$
இடைநிலை மதிப்பைவிடக் குறைவாக உள்ள மதிப்புகளின் எண்ணிக்கை	r	s	$r+s$
மொத்தம்	$p+r$	$q+s$	$N=n_1+n_2$

ஒரே இடைநிலையைப் பெற்ற முழுமைத் தொகுதிகளிலிருந்து கொடுக்கப்பட்ட கூறுகள் குழு Iஐயும், குழு IIஐயும் குறிப்பிட்டால், ஒவ்வொரு குழுவின் பாதி மதிப்புகளுள் இணைந்த இடைநிலையின் (combined median) மதிப்பை விட அதிகமாகவும், மீதிப் பாதி மதிப்புகள் இவ் விடைநிலை மதிப்பைவிடக் குறைவாகவும் இருக்கும் என்று எதிர்பார்க்கிறோம். அதாவது p, r ஆகியவைகளின் அலைவெண்கள் சமமாகவும், q, s ஆகியவைகளின் அலைவெண்கள் சமமாகவும் இருக்கும்.

குழு I, குழு II ஆகியவைகளின் அளவுகள் n_1, n_2 என்க. இரு குழுக்களின் இடைநிலையின் மதிப்பைவிட அதிகமாக உள்ள குழு I-ன் மதிப்புகளின் எண்ணிக்கை p ஆகவும், குழு II-ன் மதிப்புகள் q ஆகவும் கொண்டால், $p = \frac{n_1}{2}$, $q = \frac{n_2}{2}$ என்ற சூனிய எடுகோளின் கீழ், p, q ஆகியவைகளின் மாதிரிப் பரவலானது அதிபெருக்கப் பரவலாக (Hyper geometric) அமையும் என்பதைக் காணலாம். அதாவது,

$$P_{(p,\, q)} = \frac{\binom{p+r}{p}\binom{q+s}{q}}{\binom{n_1+n_2}{p+q}}$$ என்று அமையும்.

இரு குழுக்களிலுள்ள மொத்த வகைகளின் எண்ணிக்கையான $n_1 + n_2$ மிகச்சிறியதாக இருந்தால், H_0-ஐச் சோதனை செய்வதற்கு, ஃபிஷரின் சோதனையைப் பயன்படுத்தலாம். $n_1 + n_2$-ன் அளவு போதுமான அளவு பெரியதாக இருந்தால் 1 வரையற்ற பாகையைப் பெற்ற கைவர்க்கச் சோதனையைப் பயன்படுத்தலாம்.

$n_1 + n_2 > 40$ எனில் தொடர்ச்சிக்கான கைவர்க்கச் சோதனையைப் பயன்படுத்தவேண்டும். அதாவது 1 வரையற்ற பாகையைக் கொண்ட

$$\chi^2 = \frac{N\left(|ad-bc| - \frac{N}{2}\right)^2}{(a+b)\,(c+d)\,(a+c)\,(b+a)}$$ என்ற சூத்திரத்தைப் பயன்படுத்த வேண்டும்.

$n_1 + n_2$ ஆனது 20-க்கும் 40-க்கும் இடையே இருந்து, இணைப் பட்டியலிலுள்ள அலைவெண்ணின் மதிப்பு 5-ஐ விடக் குறைவாக இல்லாமலிருந்தால் மேலுள்ள தொடர்ச்சிக்கான கைவர்க்கச் சோதனையையே பயன்படுத்தலாம். ஆனால், இணைப்பட்டியலிலுள்ள அலைவெண்ணின் மதிப்பு 5ஐ விடக் குறைவாக இருந்தால் 'ஃபிஷரின்' சோதனையைத்தான் பயன்படுத்த வேண்டும்.

$n_1 + n_2 < 20$ எனில் 'ஃபிஷரின்' சோதனையைப் பயன்படுத்தவேண்டும்.

கவனிக்கவும்: குழு I, குழு II ஆகியவைகளின் பெரும்பாலான மதிப்புகள் இணைந்த இடைநிலை மதிப்புக்குச் சமமாக இருக்க

நேரிடலாம். அத் தருணங்களில் $n_1 + n_2$ பெரியதாக இருந்து சில வகைகள் மட்டும் இணைந்த இடைநிலை மதிப்புக்குச் சமமாக இருந்தால், அந்தச் சில வகைகளை ஆய்வினின்று புறக்கணித்துவிடலாம்; அல்லது குழுக்களை, இடைநிலையின் மதிப்புக்கு அதிகமான மதிப்பைப் பெற்ற குழு என்றும், பெறாத குழு என்றும் இரு பிரிவாகக் கொண்டு சோதனை செய்யலாம்.

குழுக்களின் மதிப்புகள் இணைந்த இடைநிலையின் மதிப்புக்குச் சமமாக இருந்தால், அவைகளின் எண்ணிக்கையை 'அதிகமான மதிப்பைப் பெறாத' என்ற வகையில் சேர்த்துக் கொள்கிறோம்.

மாதிரி: இரண்டு ரகமான மக்காச்சோளம் பயிர் செய்யப்பட்டதில் முதல் வகையில் 16 எண்ணிக்கையும் இரண்டாம் வகையில் 23 எண்ணிக்கையும் கீழ்க்கண்ட மகசூல்களை அளித்தன.

முதல் ரகம்	இரண்டாம் ரகம்
78	94
79	95
80	97
81	98
82	99
83	100
84	101
85	102
86	103
87	104
88	105
89	106
90	107
91	108
92	109
93	110
	111
	112
	113
	114
	115
	116
	117

முதல் ரகத்திற்கான இடைநிலைக்கும், இரண்டாம் ரகத்திற்கான இடைநிலைக்குமிடையே வேறுபாடு இல்லை என்ற சூனிய எடுகோளையும், முதல் ரகத்திற்கான இடைநிலை, இரண்டாம் ரகத்திற்கான இடைநிலையை விட அதிகமானது என்ற ஒருமுனைச் சோதனைக்கான மாற்று எடுகோளையும் கொண்டு மேலுள்ள விவரங்களுக்கு இடைநிலைச் சோதனை செய்வோம்.

$\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். கொடுக்கப்பட்டுள்ள விவரங்களிலிருந்து $n_1 = 16$, $n_2 = 23$ என்று அறிகிறோம். $n_1 + n_2 = 39$ ஆனது 20-க்கும் 40-க்கும் இடையே அமைகிறது. இருந்தாலும் இணைப்பட்டியலை அமைத்து, அதில் காணப்படும் அலைவெண்ணின் மதிப்புத் தெரிந்த பிறகே நாம் கைவர்க்கச் சோதனையைச் செய்யவேண்டுமா அல்லது ஃபிஷரின் சோதனையைச் செய்யவேண்டுமா என்று தீர்மானிக்க முடியும். முதல் ரகம், இரண்டாம் ரகம் ஆகியவைகளின் மதிப்புகளை இணைத்து

78 79 80 81 82 83 84 85 86 87 88 89 90
91 92 93 94 95 97 [98] 99 100 101 102 103
104 105 106 107 108 109 110 111 112 113
114 115 116 117

என்று வரிசைப்படுத்தினால் இணை இடைநிலையின் மதிப்பு 98 ஆகும். இப்பொழுது (2×2) இணைப்பட்டியலைக் கீழ்க்கண்டவாறு அமைப்போம்.

	முதல் ரகம்	இரண்டாம் ரகம்	மொத்தம்
இடைநிலை மதிப்பு 98-ஐ விட அதிகமாக உள்ள மதிப்புகளின் எண்ணிக்கை	3	15	18
இடைநிலை மதிப்பைவிட அதிகமாக இல்லா மதிப்புகளின் எண்ணிக்கை	13	8	21
மொத்தம்	16	23	39

மேலுள்ள (2×2) இணைப் பட்டியலுக்கு எதிர்பார்க்கும் அலைவெண்களை உத்தேசமாகக் கண்டால். எந்த அலைவெண்ணும்

5-க்குக் குறைவாக இல்லை என்பதாலும் $20 < n_1 + n_2 < 40$ என்பதாலும், 1 வரையற்ற பாகையைப் பெற்ற தொடர்ச்சிக்கான கைவர்க்கப் பரவலைப் பயன்படுத்துகிறோம்.

$$\chi^2 = \frac{N\left(|ad - bc| - \frac{N}{2}\right)^2}{(a+b)(c+d)(a+c)(b+d)}$$

$$= \frac{39\left(|(3)(8) - (13)(15)| - \frac{39}{2}\right)^2}{(18)(21)(16)(23)}$$

$$= \frac{39\,(151{\cdot}5)^2}{139100}$$

$$= 6{\cdot}346$$

அட்டவணை C–லிருந்து 1 வரையற்ற பாகைக்கு H_0-ன் கீழ் ஒருமுனைச் சோதனைக்கு $\chi^2 \geqslant 6{\cdot}346$-க்கான நிகழ்தகவு $p < \frac{1}{2}(0{\cdot}01)$ ஆகும். எனவே, H_0-ஐ $\alpha = 0{\cdot}01$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

மேன்-விட்னியின் U சோதனை (Mann - Whitney U Test)

இரு சார்பற்ற கூறுகள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றதா என்பதை இச் சோதனையின் மூலம் அறியலாம். இது சுட்டுறுப்பைச் சாராத முறைகளில் காணப்படும் சோதனைகளில் மிகவும் சக்திவாய்ந்த சோதனையாகும். சுட்டுறுப்பைச் சார்ந்த t-சோதனைக்கான தற்கோள்களைத் தவிர்க்க நேரிடும்போது இச் சோதனையைப் பயன்படுத்தலாம். இப்பொழுது X, Y என்ற இரு முழுமைத் தொகுதிகளைக் கருதுவோம். X-ம், Y-ம் ஒரே பரவலைப் பெற்றுள்ளன என்ற சூனிய எடுகோளையும், X-ஆனது ராண்ட முறையில் Y-ஐ விடப் பெரியது என்ற மாற்று எடுகோளையும் கொண்டு சோதனை செய்கிறோம். அதாவது x என்பது X என்ற முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட ஒரு கண்டறிந்த மதிப்பென்றும், y என்பது Y என்ற முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட ஒரு கண்டறிந்த மதிப்பென்றும் கொண்டால் X-லிருந்து வரப்பெறும் மதிப்பு, Y-லிருந்து வரப்பெறும் மதிப்பைவிடப் பெரியதாக இருப்பதற்கான நிகழ்தகவு $p(X > Y) > \frac{1}{2}$ என்பது மாற்று எடுகோளாகும்.

இச் சான்று H_1-ஐ ஆதரித்தால், பின்பு X என்ற முழுமைத் தொகுதியின் பருமன் Y என்ற முழுமைத் தொகுதியின் பருமனை

விடப் பெரியது என்று முடிவு கொள்கிறோம். இம்மாதிரியே Y-ஆனது ராண்டம் முறையில் X-ஐ விடப் பெரியதாக இருக்கும் என்ற மாற்று எடுகோளைக் கொண்டால் H_1 ஆனது $p(X > Y) < \frac{1}{2}$ என்றிருக்கும். இச் சான்று H_1-ஐ ஆதரித்தால் பின்பு Y என்ற முழுமைத் தொகுதியின் பருமனைவிடப் பெரியது என்று முடிவு கொள்கிறோம். இருமுனைச் சோதனைக்கு $p(X > Y) \neq \frac{1}{2}$ என்ற மாற்று எடுகோளைக் கொள்கிறோம்.

இரு சார்பற்ற குழுக்களில் சிறிய அளவைப் பெற்ற குழுவின் அளவு n_1 என்றும், n_2 என்பது பெரிய அளவைப் பெற்ற குழுவின் அளவு என்றும் கொள்வோம். இப்பொழுது இரு குழுக்களையும் இணைத்து ஏறு வரிசைப்படுத்திய பின்பு U என்ற புள்ளியியல் (Statistic) அளவையைக் காண்கிறோம்.

உதாரணமாக. $n_1 = 4$, $n_2 = 6$ என்ற அளவுகளைப் பெற்ற A, B என்ற இரு குழுக்களின் கண்டறிந்தவை

A	8	10	14	16		
B	5	7	9	12	13	15

என்க. இவ்விரு குழுக்களையும் இணைத்து அக் குழுக்களின் அடையாளத்தோடு, ஏறுவரிசையில் கீழ்க்கண்டவாறு அமைக்கிறோம்.

5	7	8	9	10	12	13	14	15	16
B	B	A	B	A	B	B	A	B	A

இதில் A-ன் மதிப்பானது B-ன் மதிப்பை முந்தி அமையக் கூடிய நிகழ்ச்சிகளின் மொத்த எண்ணிக்கையை U என்று குறிப்போம். அவ்வாறே, முதலில் 9-க்கு ஏற்ற Bயை முந்தி ஒரு Aயின் மதிப்பும், 12-க்கு ஏற்ற Bயை முந்தி, இரு Aயின் மதிப்புகளும் (8-க்கு ஏற்ற A, 10-க்கு ஏற்ற A), 13-க்கு ஏற்ற Bயை முந்தி இரு Aயின் மதிப்புகளும், 15-க்கு ஏற்ற Bயை முந்தி மூன்று Aயின் மதிப்புகளும் அமைந்துள்ளன. மொத்தமாகக் கொண்டால், $U = 1 + 2 + 2 + 3 = 8$ ஆகும்.

n_1 அல்லது n_2ஆனது 8-ஐ விட அதிகமில்லாமல் இருந்தால், அட்டவணை Jயைப் பயன்படுத்தி H_0-ன் கீழ் U-க்கான நிகழ்தகவை அறியலாம்.

இங்கு $n_1 = 4$, $n_2 = 6$, $U = 8$ ஆகும்.

அட்டவணை J-லிருந்து H_0-ன் கீழ் $U \leqslant 8$-க்கான நிகழ்தகவு $p = 0{\cdot}238$ ஆகும். அட்டவணை J-ல் ஒருமுனைச் சோதனைக்கான நிகழ்தகவு மதிப்புகள் கொடுக்கப்பட்டுள்ளன. இருமுனைச் சோதனைக்கு, அட்டவணையில் கொடுக்கப்பட்டுள்ள மதிப்பை இரு மடங்காகக் கொள்ள வேண்டும்.

சில சமயங்களில் நாம் காணும் U-ன் மதிப்புப் பெரியதாக இருந்தால் அட்டவணை J-ஐப் பயன்படுத்த முடியாத நிலை ஏற்படலாம். அச் சமயங்களில், அவ்வாறு காணக் கிடைக்கும் U-ஐ U' என்று குறிப்பிட்டுச் சிறிய மதிப்பைப் பெற்ற U-ஐ $U = n_1 n_2 - U'$ என்று கணக்கிடுகிறோம். இவ்வாறாக $p(U \geqslant U') = p(U \leqslant n_1 n_2 - U')$ என்பதை அறிகிறோம்.

n_1, n_2 ஆகியவைகளின் அளவுகள் ஓரளவு பெரியதாக இருந்தால் U-யைக் கணக்கிடுவது சற்றுக் கடினமாகும். ஆனால், இரு குழுக்களையும் இணைத்து ஏறு வரிசைப்படுத்தப்பட்ட தொடருக்குத் தரத்தைக் (rank) கொடுத்து,

$$U = n_1 n_2 - \frac{n_1(n_1 + 1)}{2} - R_1$$

என்ற சூத்திரத்தையோ அல்லது

$$U = n_1 n_2 - \frac{n_2(n_2 + 1)}{2} - R_2$$

என்ற சூத்திரத்தையோ பயன்படுத்தி U-ன் மதிப்பை எளிதாகக் காணலாம். இதில் R_1 என்பது n_1 அளவைப் பெற்ற $(n_1 < n_2)$ குழுவிற்கு அளிக்கப்பட்ட தரங்களின் கூட்டுத் தொகையாகும்.

R_2 என்பது n_2 அளவைப் பெற்ற $(n_2 > n_1)$ குழுவிற்கு அளிக்கப்பட்ட தரங்களின் கூட்டுத் தொகையாகும்.

மேலுள்ள சூத்திரங்களைப் பயன்படுத்தி பெறப்படும் இரு U மதிப்புகளின் சிறிய மதிப்பை U என்றும், பெரிய மதிப்பை U' என்றும் கொள்கிறோம். பின்பு அட்டவணை J-யிலிருந்து n_1, n_2, U ஆகியவைகளுக்கேற்ற p-யின் மதிப்பை அறிகிறோம். இந்த p-யின் மதிப்பு மிகைத்தன்மை மட்டமான $\alpha = 0{\cdot}05$ஐ விடக் குறைவாகவோ அல்லது சமமாகவோ இருந்தால் H_0-ஐ மறுக்கிறோம்.

மாதிரி : உயர் பட்டப்படிப்புக் கல்லூரியிலுள்ள ஒரு குறிப்பிட்ட வகுப்பில் 5 மாணவர்கள் கிராமப்புறத்திலிருந்தும்,

4 மாணவர்கள் பட்டினத்திலிருந்தும் வந்து படிப்பவர்களாக இருக்கின்றனர். இம் மாணவர்களின் உடல் கட்டமைப்பைச் சிறந்த முறையில் பெற்றவர்கள் கிராமத்திலிருந்து வந்து படிக்கும் மாணவர்களா அல்லது பட்டினத்து மாணவர்களா என்பதை அறியும் முகமாக உடல் கட்டமைப்பிற்கான மதிப்புகள் கணக்கிடப்பட்டு அட்டவணைப்படுத்தப்பட்டன.

A	B
கிராமத்து மாணவர்கள்	பட்டினத்து மாணவர்கள்
68	100
54	60
65	43
35	41
72	

இப்பொழுது கிராமத்து மாணவர்கள் சிறந்த உடல் கட்டமைப்புக்கான தகுதியைப் பட்டினத்து மாணவர்களைவிடப் பெற்றிருக்கவில்லை என்ற சூனிய எடுகோளையும், கிராமத்து மாணவர்கள் பட்டினத்து மாணவர்களைவிட அதிக உடல் கட்டமைப்புக் கொண்டுள்ளனர் என்ற மாற்று எடுகோளையும் கொண்டு, சார்பற்ற இரு மாதிரிகளுக்கான 'மேன்-விட்னி *U*' சோதனையைச் செய்வோம். n_1 என்பது இரு குழுக்களில் சிறிய அளவைப்பெற்ற குழுவின் அளவை என்பதால், $n_1 = 4, n_2 = 5$ ஆகும். $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கொள்வோம். இரு குழுக்களையும் இணைத்து ஏறுவரிசையில், ஒவ்வொரு மதிப்பும் எக்குழுவைச் சேர்ந்தது என்பதைக் குறுப்பிட்டு அமைப்போம். அவ்வாறாக,

35	41	43	54	60	65	68	72	100
A	*B*	*B*	*A*	*B*	*A*	*A*	*A*	*B*

என்று வரிசைப் படுத்துகிறோம். இவைகளுக்குத் தரத்தைக் கொடுத்தால்,

மதிப்பு ;	35	41	43	54	60	65	68	72	100
குழு	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*A*	*B*
தரம்	1	2	3	4	5	6	7	8	9

இதில் $n_1 = 4$ என்ற அளவைப் பெற்ற குழு B-க்கான தரங்களின் மொத்தம் $R_1 = 2+3+5+9 = 19$ ஆகும். இதே போன்று $n_2 = 5$ என்ற அளவைப் பெற்ற குழு A-க்கான தரங்களின் மொத்தம்

$$R_2 = 1+4+6+7+8 = 26 \text{ ஆகும்.}$$

$$\text{இப்பொழுது } U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= 4 \times 5 + \frac{4(4+1)}{2} - 19$$

$$= 11 \text{ ஆகும்.}$$

$$\text{அல்லது } U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

$$= 4 \times 5 + \frac{5(5+1)}{2} - 26$$

$$= 9 \text{ ஆகும்.}$$

இதில் சிறிய U-ன் மதிப்பான 9-ஐத்தான் நாம் கருத்தில் கொள்கிறோம். U-ன் பெரிய மதிப்பான 11-ஐ U' என்று குறிப்பிடுகிறோம்.

எனவே $U = n_1 n_2 - U' = 4 \times 5 - 11 = 9$ ஆகும். இப்பொழுது H_0-ன் கீழ் அட்டவணை J-லிருந்து $n_1 = 4$, $n_2 = 5$ $U \leqslant 9$-க்கான நிகழ்தகவு $p = 0{\cdot}452$ என அறிகிறோம். கொடுக்கப்பட்டுள்ள விவரமானது $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் H_0-ஐ மறுப்பதற்கான தடயங்களைப் பெற்றிருக்கவில்லை என்று தீர்மானிக்கிறோம்.

குறிப்பு: n_2-ன் அளவானது 9-லிருந்து 20 வரையிருந்தால் அட்டவணை J-யைப் பயன்படுத்த முடியாது. அச் சமயங்களில் அட்டவணை K-யைப் பயன்படுத்தி $\alpha = 0{\cdot}001$, $0{\cdot}01$, $0{\cdot}025$, $0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்திற்கு U-க்கான தீர்வு கட்டமான மதிப்புகளைப் பெறலாம். இருமுனைச் சோதனை என்றால் $0{\cdot}002$, $0{\cdot}02$, $0{\cdot}05$, $0{\cdot}10$ என்ற மிகைத்தன்மை மட்டங்களுக்குத் தீர்வு கட்டமான மதிப்புகளை அறியலாம்.

அட்டவணை K-ல் U-ன் தீர்வு கட்டமான மதிப்புகள் மட்டுமே கொடுக்கப்பட்டுள்ளன. இது சரியான நிகழ்தகவு மதிப்புகளைக்

கொடுக்கவில்லை. $n_1 \leqslant 20$, $9 < n_2 < 20$ ஆகியவைகளுக் கேற்ற U-ன் கண்டறிந்த மதிப்பானது α என்ற மிகைத்தன்மை மட்டத்திற்கு அட்டவணை K-ல் கொடுக்கப்பட்டுள்ள மதிப்பை விடக்குறைவாகவோ அல்லது சமமாகவோ இருந்தால் H_0-ஐ அம்மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

உதாரணமாக $n_1 = 6$, $n_2 = 16$, $U = 16$ எனில் $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தில் அட்டவணை K-ல் காணும் மதிப்பு 12 ஆகும்.

எனவே, கண்டறிந்த U-ன் மதிப்பான 16, அட்டவணை K-ன் மதிப்பான 12-யை விட அதிகமாக உள்ளதால், H_0-ஐ ஏற்றுக் கொள்கிறோம்.

இருமுனைச் சோதனை என்றால் $\alpha = 0{\cdot}02$ என்று கொண்டு H_0-ஐ ஏற்றுக்கொள்கிறோம்.

$n_2 > 20$ என்றால் அக் கூறுகளைப் பெரிய கூறுகள் என்று கூறு கிறோம். அப்பொழுது J, K ஆகிய அட்டவணைகளைப் பயன்படுத்த முடியாது. எனினும் n_1, n_2 பெரியதாக இருந்தால் U-ன் மாதிரிப் பரவலானது இயல்நிலைப் பரவலை அணுகுகிறது என்ற முடிவைப் பயன்படுத்தி,

$$z = \frac{U-\mu}{\sigma}$$

என்ற புள்ளியியல் அளவைக் காண்கிறோம்.

இதில் $\mu = \dfrac{n_1 n_2}{2}$, $\sigma = \sqrt{\dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$ ஆகும்.

z ஆனது பூச்சியம் சராசரி மதிப்பையும், 1 மாறுபாட்டையும் பெற்ற இயல்நிலைப் பரவலாக அமைகிறது. எனவே, H_0-ன் கீழ் நிகழ்ச்சி நிகழ்வதற்கான நிகழ்தகவு மதிப்பை அட்டவணை A-லிருந்து பெற்று, சூனிய எடுகோளை ஏற்றுக்கொள்ளவோ அல்லது மறுக்கவோ செய்யலாம். z-ன் மட்டு மதிப்பைக் கொள்வ தால் U-ன் மாதிரிப் பரவலுக்கு இயல்நிலைத் தோராயத்தைப் பயன்படுத்தும்போது, U-ஐக் காண உதவும் இரு சூத்திரங்களில் எவையேனும் ஒன்றைப் பயன்படுத்தலாம்.

மாதிரி : ஒரு வகுப்பிலுள்ள மாணவர்கள் ராண்டம் முறையில் இரு பிரிவுகளாகப் பிரிக்கப்பட்டனர். ஒரு குழுவிலுள்ள மாணவர்களுக்கு ஆசிரியரின் சொற்படி சீரான முறையில் படிக்கக்

கற்றுக்கொடுக்கப்பட்டது. மற்ெருரு குழுவிலுள்ள மாணவர் களுக்கு அவர்கள் விரும்பிய முறையில் ஓர் ஆசிரியரின் மேற் பார்வையில் படிக்கக் கற்றுக்கொண்டனர். இவ்விரு குழுவிலுள்ள மாணவர்களுக்கு வருட இறுதியில் படிக்கும் திறனுக்கான தேர்வு நடத்தப்பட்டு அதற்கான மதிப்புகள் கீழ்க்கண்டவாறு அட்ட வணைப் படுத்தப்பட்டன.

முதற்குழு	இரண்டாம் குழு
23	27
22	26
22	25
20	25
20	25
20	24
20	24
19	24
18	23
18	23
17	23
17	22
17	22
17	22
17	22
16	21
	21
	20
	20
	20
	18
	18
	18

மேன்-விட்னி *U* சோதனையைப் பயன்படுத்தி இவ்விரு சார்பற்ற குழுக்களின் சராசரிகளுக்குள் வேறுபாடு உள்ளதா என்பதை அறிவோம்.

இரு குழுக்களின் சராசரியும் சமமானதே என்ற சூனிய எடு கோளையும், முதற்குழுவின் சராசரி இரண்டாம் குழுவின் சராசரியைவிடப் பெரியது என்ற மாற்று எடுகோளையும் கொள்வோம். $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம்.

இங்கு $n_1 = 16$, $n_2 = 23$ ஆகும். $n_2 > 20$

என்பதால், $$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 \; n_2 \; (n_1 + n_2 + 1)}{12}}} \cap N(0, 1)$$

என்ற புள்ளியியல் அளவையின் மதிப்பைக் காண வேண்டும். இதில் $U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$ ஆகும். R_1-ஐக் காண, இரு குழுவையும் இணைத்து வரிசைப்படுத்திய பின்பு, தரங்களைக் கொடுக்கவேண்டும். பிறகு n_1 அளவைப் பெற்ற குழுவிலிருந்து வரப்பெற்ற மதிப்புகளுக்கான தரங்களைக் கூட்டவேண்டும். இதனைக் கீழ்க்காணும் அட்டவணையில் காணலாம்.

முதற்குழு $n_1 = 16$	தரம்	இரண்டாம் குழு $n_2 = 23$	தரம்
23	29·5	27	39
22	24·5	26	38
22	24·5	25	36
20	16	25	36
20	16	25	36
20	16	24	33
20	16	24	33
19	12	24	33
18	9·5	23	29·5
18	9·5	23	29·5
17	5	23	29·5
17	5	22	24·5
17	5	22	24·5
17	5	22	24·5
17	5	22	24.5
16	1·5	21	20·5
		21	20·5
	$R_1 = 200·0$	20	16
		20	16
		20	16
		18	9·5
		18	9·5
		18	9·5
			$R_2 = 580·0$

$R_1 = 200$ ஆகும். எனவே

$$U = 16 \times 23 + \frac{16(16+1)}{2} - 200$$

$$= 304.$$

$$\text{எனவே } z = \frac{304 - \frac{16(23)}{2}}{\sqrt{\frac{16(23)(16+23+1)}{12}}}$$

$$= 3{\cdot}43 \text{ ஆகும்.}$$

அட்டவணை A-லிருந்து $z \geqslant 3{\cdot}43$ ஆனது H_0-ன் கீழ் $p < 0{\cdot}0003$ என்ற ஒருமுனை நிகழ்தகவைப் பெற்றுள்ளது. இந்த p–ன் மதிப்பு $\alpha = 0{\cdot}01$-ஐ விடச் சிறியதாக இருப்பதால் H_0-ஐ மறுக்கிறோம். அதாவது இரு குழுக்களின் சராசரிகளும் சமமானதல்ல என்றோ, அல்லது இரு குழுக்களின் சராசரிகளுக்கிடையே வேறுபாடு உள்ளது என்றோ $\alpha = 0{\cdot}01$ மிகைத்தன்மை மட்டத்தில் உய்த்துணர்கிறோம்.

மேலுள்ள மாதிரிக் கணக்கில் முதற்குழுவிலும், இரண்டாவது குழுவிலும் ஒத்த மதிப்புடைய கண்டறிந்தவை இருந்தன. இவைகளுக்குத் தரத்தைக் கொடுக்கும்போது, தரங்கள் சமமாகப் பங்கிடப்பட்டது. இம் மாதிரியான ஒத்தவை (ties) இருப்பின் மிகத்துல்லியமான z-ன் மதிப்பைக் கீழ்க்காணும் சூத்திரம் மூலம் காண்கிறோம்.

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)}\left\{\frac{(n_1+n_2)^3-(n_1+n_2)}{12} - \Sigma T\right\}}}$$

$$\text{இதில் } U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \text{ ஆகும்.}$$

$$T = \frac{t^3 - t}{12} \text{ ஆகும்.}$$

t-ஆனது ஒரே தரத்திற்கு எத்தனை கண்டறிந்தவை இருக்கின்றன என்பதைக் குறிக்கிறது.

அவ்வாறே மேலுள்ள மாதிரிக் கணக்கில்,

16 ஆனது 2 முறையும்
17 ஆனது 5 முறையும்
18 ஆனது 4 முறையும்
20 ஆனது 7 முறையும்
21 ஆனது 2 முறையும்
22 ஆனது 6 முறையும்
23 ஆனது 4 முறையும்
24 ஆனது 3 முறையும்
25 ஆனது 3 முறையும் நிகழ்ந்துள்ளன.

எனவே t : 2, 5, 4, 7, 2, 6, 4, 3, 3 ஆகும்.

$$\Sigma T = \Sigma \left(\frac{t^3 - t}{12}\right)$$

$$= \frac{2^3 - 2}{12} + \frac{5^3 - 5}{12} + \frac{4^3 - 4}{12} + \frac{7^3 - 7}{12}$$

$$+ \frac{2^3 - 2}{12} + \frac{6^3 - 6}{12} + \frac{4^3 - 4}{12} + \frac{3^3 - 3}{12} + \frac{3^3 - 3}{12}$$

$$= 0{\cdot}5 + 10{\cdot}0 + 5{\cdot}0 + 28{\cdot}0 + 0{\cdot}5 + 17{\cdot}5$$

$$+ 5{\cdot}0 + 2{\cdot}0 + 2{\cdot}0$$

$$= 70{\cdot}5$$ ஆகும்.

எனவே,

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\left\{\frac{n_1 \; n_2}{(n_1+n_2)(n_1+n_2-1)}\right\} \left\{\frac{(n_1+n_2)^3 - (n_1+n_2)}{12} - \Sigma T\right\}}}$$

$$= \frac{304 - \frac{(16)\,(23)}{2}}{\sqrt{\left\{\frac{(16)\,(23)}{36\,(39-1)}\right\} \left\{\left(\frac{(39)^3 - 39}{12}\right) - 70{\cdot}5\right\}}}$$

$$= \frac{304 - 184}{\sqrt{\left(\frac{46}{171}\right) \left(4816{\cdot}5 - 70{\cdot}5\right)}}$$

$$= \frac{120}{\sqrt{\left(\frac{46}{171}\right)\left(4746 \cdot 0\right)}}$$

$$= \frac{120}{\sqrt{1274 \cdot 2}} = \frac{120}{35 \cdot 6}$$

$= 3{\cdot}45$ ஆகும்.

ஆனால், முன்பு கண்ட z-ன் மதிப்பு $3{\cdot}43$ ஆகும். இவைகளுக்குள்ள வித்தியாசம் சிறிதே எனினும் ஒத்தவை நிகழும்போது, துல்லியமான z-ன் மதிப்பை இச் சூத்திரத்தைப் பயன்படுத்திக் காணுதல் சாலச் சிறந்ததாகும். $z > 3{\cdot}45$-க்கு, ஒருமுனைச் சோதனைக்கு ஏற்ற நிகழ்தகவின் மதிப்பு $p \leqslant 0{\cdot}0003$ என்று அறிகிறோம்.

'கோல்மோகோராவ்–ஸ்மிர்னவி'ன் இருகூறுச் சோதனை (Kolmogorov-Smirnov Two Sample Test)

இரு வெவ்வேறு முழுமைத் தொகுதியிலிருந்து எடுக்கப்படும் இரு மாதிரிகளின் பரவல் சார்புகள் ஒத்தவைகளா என்பதை அறிய இச் சோதனை பயன்படுகிறது. மற்றச் சோதனைகளான இடைநிலைச் சோதனை, மேன்-விட்னி சோதனை அல்லது சுட்டுறுப்பைச் சார்ந்த 't' சோதனை ஆகியவைகளும் பொருத்தமானவை என்றாலும் இவை இரு சராசரி, இரு இடைநிலை ஆகியவைகளிடையே உள்ள வேறுபாடுகளை அறிய துல்லியமான சோதனைகளாக இருக்கலாம். ஆனால், மாறுபாடுகளுக்கிடையே (variances) உள்ள வேறுபாடுகளை அறிவதற்கு இச் சோதனைகள் ஏற்றவைகளாக இருக்கமாட்டா.

$X_1, X_2, \ldots\ldots X_n$; $Y_1, Y_2 \ldots\ldots Y_m$ என்ற n, m அளவுகளை முறையே பெற்ற இரு சார்பற்ற ராண்டம் மாதிரிகளைக் கருதுக.

$F(x)$, $G(x)$ என்பவை இவ்விரு ராண்டம் மாறிகளின் முறையான தெரியாத (unknown) பரவல் சார்புகள் என்க. இவ்விரு கூறுகளும் ஒன்றை ஒன்று சார்பற்றன என்றும், ராண்டம் மாறிகள் தொடர்ச்சியானவை என்றும் கொள்கிறோம். ராண்டம் மாறிகள் தொடர்ச்சியற்றனவாக இருந்தாலும் இச் சோதனை ஏற்புடையதாகும்.

இருமுனைச் சோதனைக்கு $-\infty$லிருந்து $+\infty$வரை எல்லா x-களுக்கும், $F(x) = G(x)$ என்ற சூனிய எடுகோளையும் குறைந்

தது ஒரு x-க்காவது $F(x) \neq G(x)$ என்ற மாற்று எடுகோளையும் கொள்கிறோம்.

ஒருமுனைச் சோதனைக்கு, $-\infty$லிருந்து, $+\infty$வரை எல்லா x-களுக்கும் $F(x) \leqslant G(x)$ என்ற சூனிய எடுகோளையும், குறைந்தது ஒரு x-க்காவது $F(x) > G(x)$ என்ற மாற்று எடுகோளையும் கொள்கிறோம்.

மற்றொரு விதமாக ஒருமுனைச் சோதனைக்கு $-\infty$லிருந்து $+\infty$வரை எல்லா x-களுக்கும் $F(x) \geqslant G(x)$ என்ற சூனிய எடுகோளையும், குறைந்தது ஒரு x-க்காவது $F(x) < G(x)$ என்ற மாற்று எடுகொளையும் கொள்கிறோம்.

$X_1, X_2, \ldots, X_n$ என்ற ராண்டம் மாதிரியை அடிப்படையாகக் கொண்ட எம்பெரிகல் பரவல் சார்பு $S_1(x)$ என்றும், $Y_1, Y_2, \ldots, Y_m$ என்ற மற்றொரு ராண்டம் மாதிரியை அடிப்படையாகக் கொண்ட பரவல் சார்பு $S_2(x)$ என்றும் கொள்க. இப்பொழுது முன் குறிப்பிட்ட இருமுனைச் சோதனையான $-\infty$-ல் இருந்து $+\infty$ வரை எல்லா x-க்கும் $F(x) = G(x)$ என்ற சூனிய எடுகோளையும் குறைந்தது ஒரு x-க்காவது $F(x) \neq G(x)$ என்ற மாற்று எடுகோளையும் கருதினால், T_1 என்ற புள்ளியியல் அளவையை,

$$T_1 = \sup_x \left| S_1(x) - S_2(x) \right|$$

என்று குறிப்பிடுகிறோம்.

ஒருமுனைச் சோதனையான $-\infty$லிருந்து, $+\infty$ வரை எல்லா x-க்கும், $F(x) \leqslant G(x)$ என்ற சூனிய எடுகோளையும் குறைந்தது ஒரு x-க்காவது $F(x) > G(x)$ என்ற மாற்று எடுகோளையும் கருதும் பொழுது, T_1^+ என்ற புள்ளியியல் அளவையை,

$$T_1^+ = \sup_x \left[S_1(x) - S_2(x) \right]$$

என்று குறுப்பிடுகிறோம். இதே போன்று மற்றோர் ஒருமுனைச் சோதனையான $-\infty$ லிருந்து $+\infty$ வரை எல்லா x-க்கும் $F(x) \geqslant G(x)$ என்ற சூனிய எடுகோளையும் குறைந்தது ஒரு x-க்காவது $F(x) < G(x)$ என்ற மாற்று எடுகோளையும் காணும்போது T_1^- என்ற புள்ளியியல் அளவையை,

$$T_1^- = \sup_x \left[S_2(x) - S_1(x) \right]$$

என்றும் குறிப்பிடுகின்றோம்.

தகுந்த புள்ளியியல் அளவையான T_1 அல்லது T_1^+ அல்லது T_1^- ஆனது $n = m$ ஆக இருந்தால் அட்டவணை L-ல் உள்ள $(1-\alpha)$ பகுதிமானத்திற்கு ஏற்ற மதிப்பைவிட அதிகமாக இருந்தாலோ அல்லது $n \neq m$ எனில் அட்டவணை M-ல் உள்ள $(1-\alpha)$ பகுதிமானத்திற்கேற்ற மதிப்பைவிட அதிகமாக இருந்தாலோ, α என்ற மிகைத்தன்மை மட்டத்தில் H_0-ஐ மறுக்கிறோம்.

ஒருமுனைச் சோதனைக்கு அட்டவணையில், ·90, ·95, ·975, ·99, ·995 ஆகிய பகுதிமானங்களுக்கு அட்டவணைப்படுத்தப்பட்டுள்ளது. இருமுனைச் சோதனைக்கு 0·80, 0·90, 0·95, 0·98, 0·99 ஆகிய பகுதிமானங்களுக்கு அட்டவணைப்படுத்தப்பட்டுள்ளது.

பெரிய கூறுகளுக்கு அட்டவணைகளின் கீழ்ப்பகுதியில் தோராய மதிப்புகள் கொடுக்கப்பட்டுள்ளன.

மாதிரி: X_1, X_2, ... X_9 என்ற மாணவர்களை A என்ற ஒரு பட்டினத்தின் ஒரு பகுதியிலிருந்து ராண்டம் முறையிலும் தேர்ந்தெடுக்கப்பட்டுக் கல்வியறிவுச் சோதனை செய்யப்பட்டது. இதே போன்று Y_1, Y_2 ...... Y_{15} என்ற 15 மாணவர்களை B என்ற மற்றொரு பட்டினத்தின் ஒரு பகுதியிலிருந்து ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டு அதே கல்வியறிவுச் சோதனை செய்யப்பட்டது. இவ்விரு குழுக்களும் பெற்ற மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டன. இரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட இவ்விரு கூறுகளிடையே கல்வியறிவில் வித்தியாசம் உள்ளதா என்பதை 'கோல்மோகோராவ்-ஸ்மிர்னுவ்' சோதனையைப் பயன்படுத்திக் காண்போம்.

பட்டினம் A	பட்டினம் B
X_i	Y_i
17.6	15·2
18·4	15·7
18·6	15·9
18·7	16·5
19·3	16·8
19·9	18·2
20·1	19·1
20·6	19·8
21·2	20·8
	21·3
	21·5
	22·3
	22·5
	23·4
	24·6

இப்பொழுது $-\infty$ லிருந்து $+\infty$ வரை, எல்லா x-க்கும் இரு முழுமைத் தொகுதியும் ஒரே பரவல் சார்பைப் பெற்றுள்ளன என்ற சூனிய எடுகோளையும், குறைந்தது ஒரு x-க்காவது இரு பரவல் சார்புகளும் வேறுபட்டுள்ளன என்ற மாற்று எடுகோளையும் கொள்வோம். அதாவது $-\infty$ லிருந்து, $+\infty$ வரை எல்லா x-க்கும் $F(x) = G(x)$ என்ற சூனிய எடுகோளையும், குறைந்தது ஒரு x-க்காவது $F(x) \neq G(x)$ என்ற மாற்று எடுகோளையும் கொள்வோம். இதில் $F(x)$ என்பது, முதல் முழுமைத் தொகுதியின் பரவல் சார்பாகும். $G(x)$ என்பது இரண்டாம் முழுமைத் தொகுதியின் பரவல் சார்பாகும். இப்பொழுது இவ்விரு கூறுகளையும் சிறியது முதல், பெரியதுவரை வரிசைப்படுத்தியும், மேலும் இவ்விரு கூறுகளின் மதிப்புகளுக்கேற்ற, அனுபவப் பரவல் (emperical distribution) சார்புகளுக்கிடையேயுள்ள வேறுபாட்டையும், கீழ்க்கண்டவாறு அட்டவணைப்படுத்துகிறோம்.

இங்கு $n_1 = 9$, $n_2 = 15$ ஆகும்.

X_i	Y_i	$S_1(x) - S_2(x)$
	15·2	$0 - \frac{1}{15} = -\frac{1}{15}$
	15·7	$0 - \frac{2}{15} = -\frac{2}{15}$
	15·9	$0 - \frac{3}{15} = -\frac{3}{15}$
	16·5	$0 - \frac{4}{15} = -\frac{4}{15}$
	16·8	$0 - \frac{5}{15} = -\frac{5}{15}$
17·6		$\frac{1}{9} - \frac{5}{15} = -\frac{2}{9}$
	18·2	$\frac{1}{9} - \frac{6}{15} = -\frac{13}{45}$
18·4		$\frac{2}{9} - \frac{6}{15} = -\frac{8}{45}$
18·6		$\frac{3}{9} - \frac{6}{15} = -\frac{1}{15}$
18·7		$\frac{4}{9} - \frac{6}{15} = \frac{2}{45}$

8

X_i	Y_i	$S_1(x) - S_2(x)$
	19·1	$\frac{4}{9} - \frac{7}{15} = -\frac{1}{45}$
19·3		$\frac{5}{9} - \frac{7}{15} = \frac{4}{45}$
	19·8	$\frac{5}{9} - \frac{8}{15} = \frac{1}{45}$
19·9		$\frac{6}{9} - \frac{8}{15} = \frac{2}{15}$
20·1		$\frac{7}{9} - \frac{8}{15} = \frac{11}{45}$
20·6		$\frac{8}{9} - \frac{8}{15} = \frac{16}{45}$
	20·8	$\frac{8}{9} - \frac{9}{15} = \frac{13}{45}$
21·2		$1 - \frac{9}{15} = \frac{2}{5}$
	21·3	$1 - \frac{10}{15} = \frac{1}{3}$
	21·5	$1 - \frac{11}{15} = \frac{4}{15}$
	22·3	$1 - \frac{12}{15} = \frac{1}{5}$
	22·5	$1 - \frac{13}{15} = \frac{2}{15}$
	23·4	$1 - \frac{14}{15} = \frac{1}{15}$
	24·6	$1 - \frac{15}{15} = 0$

இந்த இருமுனைச் சோதனைக்கு,

$$T_1 = \sup_x \left| S_1(x) - S_2(x) \right|$$

$$= \frac{2}{5} = 0 \cdot 4 \text{ ஆகும்.}$$

அதாவது மிகப் பெரிய $[S_1(x) - S_2(x)]$ மட்டு மதிப்பு (absolute value) $\frac{2}{5} = 0\cdot 4$ ஆகும்.

இப்பொழுது $\alpha = 0\cdot 05$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதினுல், $n_1 \neq n_2$ என்பதால் அட்டவணை M-லிருந்து, $(1 - \alpha) = 1 - 0\cdot 05 = 0\cdot 95$ பகுதி மானத்திற்கு, $n = 9$, $m = 15$-க் கேற்ற $w_{\cdot 95}$ன் மதிப்பு $\frac{8}{15}$ ஆகும். அதாவது, $\frac{8}{15} = 0\cdot 533$ ஆகும். T_1-ன் மதிப்பான $0\cdot 400$ ஆனது $0\cdot 533$ஐ விடக் குறைவாக உள்ளதால் H_0ஐ விட $0\cdot 05$ மிகைத்தன்மை மட்டத்தில் ஏற்றுக் கொள்கிறேம். அதாவது, இரு பட்டின மாணவர்களின் கல்வித் திறன் சமமானது என்பதை ஏற்றுக்கொள்கிறேம். அட்டவணையில் கண்ட $0\cdot 95$ பகுதிமான மதிப்பான $0\cdot 53$ஐ அணுகுகோடு பரவலைத் தழுவிக் கீழ்க்கண்ட சூத்திரம் மூலமாகத் தோராயமாகப் பெறலாம்.

$$w_{\cdot 95} \simeq 1\cdot 36 \sqrt{\frac{m + n}{mn}} = 0\cdot 573 \text{ ஆகும்.}$$

ஆனல், இது $0\cdot 533$ஐ விடச் சிறிது அதிகமாக உள்ளது என்பதைத் தவிர மற்றபடி தோராய மதிப்பை அட்டவணையின்றிப் பெறலாம்.

இருமுனைச் சோதனையைக் கருதுவதற்குப் பதில் ஒருமுனைச் சோதனையைக் கருதியிருந்தால்,

$$T_1^+ = \sup_x \left[S_1(x) - S_2(x) \right] = \frac{2}{5} = 0\cdot 4 \text{ ஆகும்.}$$

இவ்வாறே,

$$T_1^- = \sup_x \left[S_2(x) - S_1(x) \right] = \frac{1}{3} = 0\cdot 333 \text{ ஆகும்.}$$

ஒருமுனைச் சோதனைக்கு அட்டவணையிலிருந்து, $1 - \alpha = 1 - 0\cdot 05 = 0\cdot 95$ பகுதிமானத்திற்கு $w_{\cdot 95}$ன் மதிப்பு $\frac{22}{45} = 0\cdot 488$ ஆகும். கண்டறிந்த T_1^+, T_1^- ஆகிய மதிப்புகள் அட்டவணை மதிப்பான $0\cdot 488$ஐ விடக் குறைவாக உள்ளதால் H_0ஐ $\alpha = 0\cdot 05$ மிகைத்தன்மை மட்டத்தில் ஏற்றுக் கொள்கிறேம்.

குறிப்பு : ஒருமுனைச் சோதனைக்குப் பெரிய கூறுகளாக இருந்தால்,

$$T^{+} = \sup_{x} \; [S_1(x) - S_2(x)] \text{ ஐக் கண்டபின்}$$

$$\chi^2 = 4\,(T^{+})^2 \frac{n\,m}{n+m} \text{ ஆனது}$$

2 வரையற்ற பாகையோடு கைவர்க்கப் பரவலாக அமைந்துள்ளது என்று குட்மேன் (1954) என்பவர் நிரூபித்துள்ளார்.

இவ்வாறாக மேலுள்ள மாதிரிக் கணக்கில்,

$$T_1^{+} = 0{\cdot}400 \text{ ஆகும்.}$$

$$\text{இவ்வாறு, } \chi^2 = 4\,(\cdot 400)^2 \frac{(9)(15)}{24}$$

$$= 4\,(\cdot 16)\frac{45}{8}$$

$$= 0{\cdot}36$$

அட்டவணை C-லிருந்து χ^2-ன் 2 வரையற்ற பாகைகளுக்கு $\alpha = 0{\cdot}01$ மிகைத்தன்மை மட்டத்தின் மதிப்பு $9{\cdot}21$ ஆகும். கண்டறிந்த χ^2-ன் மதிப்பு அட்டவணையின் மதிப்பைவிட அதிகமாக உள்ளதால் H_0-ஐ $\alpha = 0{\cdot}01$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

சிறிய கூறுகளாக இருந்தாலும், இந்த கைவர்க்கத் தோராயத்தைப் பயன்படுத்தலாம். இதனால், ஏற்படக்கூடிய பிழை மிகச் சொற்பமானதாக இருக்கும் என்பதால் உய்த்துணர்வு எவ்விதத்திலும் பாதிக்கப்படுவதில்லை.

'க்ராமர்-வான் மைஸஸி'ன் இருகூறுச் சோதனை (Cramer-von Mises Two Sample Test)

$F(x)$, $G(x)$ என்ற தெரியாத பரவல் சார்புகளை முறையே பெற்ற $X_1, X_2, \ldots\ldots, X_n; Y_1, Y_2, \ldots\ldots Y_n$ என்ற இரு சார்பற்ற ராண்டம் கூறுகளைக் கருதுக. இந்த ராண்டம் மாறிகள் தொடர்ச்சியானது என்றும், இரு ராண்டம் மாதிரிகளும் ஒன்றுக்கொன்று சார்பற்றன என்றும் கொள்கிறோம்.

$-\infty$ யிலிருந்து $+\infty$ வரை எல்லா x-க்கும் $F(x) = G(x)$ என்ற சூனிய எடுகோளையும், குறைந்தது ஒரு x-க்காவது

$F(x) \neq G(x)$ என்ற மாற்று எடுகோளையும் கருதி இச் சோதனையைச் செய்கிறோம். இரு கூறுகளின் எம்பெரிகல் பரவல் சார்புகளை முறையே $S_1(x)$, $S_2(x)$ ஏற்றுக்கொண்டு T_2 என்ற புள்ளியியல் அளவையை,

$$T_2 = \frac{mn}{(m+n)^2}\left\{ \sum_{i=1}^{n} \left[S_1(x_i) - S_2(x_i) \right]^2 + \sum_{j=1}^{m} \left[S_1(y_j) - S_2(y_j) \right] \right\}^2$$

என்று வரையறை செய்கிறோம்.

ஒரு குறிப்பிட்ட α என்ற மிகைத்தன்மை மட்டத்திற்கு $(1-\alpha)$ ன் பகுதிமானமான $w_{1-\alpha}$ ன் மதிப்பைவிட T_2-ன் மதிப்பு அதிகமாக இருந்தால் H_0-ஐ மறுக்கிறோம்.

அணுகுகோடு பரவலை அடிப்படையாகக் கொண்டு m, n களின் பெரிய அளவுகளுக்கு $w_{1-\alpha}$ என்ற பகுதிமானங்களின் மதிப்புகள் பின்வருமாறு :

$w_{.10}$	=	0·046	$w_{.50}$	=	0·119	$w_{.90}$	=	0·347
$w_{.20}$	=	0·062	$w_{.60}$	=	0·147	$w_{.95}$	=	0.461
$w_{.30}$	=	0·079	$w_{.70}$	=	0·184	$w_{.99}$	=	0.743
$w_{.40}$	=	0.097	$w_{.80}$	=	0·241	$w_{.999}$	=	1·168

இம் மதிப்புகளை ஆண்டர்ஸன், டார்லிங் (1952) ஆகியவர்கள் கணக்கிட்டனர். n, m ஆகியவைகளின் அளவுகள் ஓரளவு சிறியதாக இருந்தாலும் இந்தப் பகுதிமான மதிப்புகளைப் பயன்படுத்தலாம்.

மாதிரி : மாதிரி 2-ல் உள்ள விவரங்களைக் கொண்டு புள்ளியியல் அளவையான T_2 ஐக் காண்போம்.

$$T_2 = \frac{mn}{(m+n)^2}\left\{ \left(\sum_{i=1}^{n} \left[S_1(x_i) - S_2(x_i) \right]^2 \right) + \sum_{j=1}^{m} \left[S_1(y_j) - S_2(y_j) \right]^2 \right\}$$

இப்பொழுது, $\sum_{i=1}^{9}\left[S_1(x_i) - S_2(x_i)\right]^2$

$$= \left\{\left(\frac{1}{9}-\frac{5}{15}\right)^2+\left(\frac{2}{9}-\frac{6}{15}\right)^2+\left(\frac{3}{9}-\frac{6}{15}\right)^2\right.$$

$$+\left(\frac{4}{9}-\frac{6}{15}\right)^2+\left(\frac{5}{9}-\frac{7}{15}\right)^2+\left(\frac{6}{9}-\frac{8}{15}\right)^2$$

$$\left.+\left(\frac{7}{9}-\frac{8}{15}\right)^2+\left(\frac{8}{9}-\frac{9}{15}\right)^2+\left(1-\frac{9}{15}\right)^2\right\}$$

$$= \left\{\left(-\frac{2}{9}\right)^2+\left(-\frac{8}{45}\right)^2+\left(-\frac{1}{15}\right)^2+\left(\frac{2}{45}\right)^2\right.$$

$$\left.+\left(+\frac{4}{45}\right)^2+\left(\frac{2}{15}\right)^2+\left(\frac{11}{45}\right)^2+\left(\frac{16}{45}\right)^2+\left(\frac{2}{5}\right)^2\right\}$$

$$= \left\{\frac{4}{81}+\frac{64}{2025}+\frac{1}{225}+\frac{4}{2025}+\frac{16}{2025}+\frac{4}{2025}\right.$$

$$\left.+\frac{121}{2025}+\frac{256}{2025}+\frac{4}{25}\right\}$$

$$= 0{\cdot}459$$

மேலும் $\sum_{j=1}^{15}\left[S_1(y_j) - S_2(y_j)\right]^2$

$$= \left(0-\frac{1}{15}\right)^2+\left(0-\frac{2}{15}\right)^2+\left(0-\frac{3}{15}\right)^2$$

$$+\left(0-\frac{4}{15}\right)^2+\left(0-\frac{5}{15}\right)^2+\left(\frac{1}{9}-\frac{6}{15}\right)^2$$

$$+\left(\frac{4}{9}-\frac{7}{15}\right)^2+\left(\frac{5}{9}-\frac{8}{15}\right)^2+\left(\frac{8}{9}-\frac{9}{15}\right)^2$$

$$+\left(1-\frac{10}{15}\right)^2+\left(1-\frac{11}{15}\right)^2+\left(1-\frac{12}{15}\right)^2$$

$$+\left(1-\frac{13}{15}\right)^2+\left(1-\frac{14}{15}\right)^2+\left(1-1\right)^2$$

$$= \left\{ \left(-\frac{1}{15}\right)^2 + \left(-\frac{2}{15}\right)^2 + \left(-\frac{3}{15}\right)^2 + \left(-\frac{4}{15}\right)^2 + \left(-\frac{5}{16}\right)^2 + \left(-\frac{13}{45}\right)^2 + \left(-\frac{1}{45}\right)^2 + \left(\frac{1}{45}\right)^2 + \left(\frac{13}{45}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{4}{15}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{2}{15}\right)^2 + \left(\frac{1}{15}\right) \right\}$$

$$= \left\{ \frac{1}{225} + \frac{4}{225} + \frac{9}{225} + \frac{16}{225} + \frac{25}{225} + \frac{169}{225} + \frac{1}{2025} + \frac{1}{2025} + \frac{169}{2025} + \frac{1}{9} + \frac{16}{225} + \frac{1}{25} + \frac{4}{225} + \frac{1}{225} \right\}$$

$$+ \frac{1330}{2025} = 0{\cdot}657$$ ஆகும்.

• எனவே,

$$T_2 = \frac{mn}{(m+n)^2} \left\{ \sum_{i=1}^{n} \left[S_1(x_i) - S_2(x_i) \right]^2 + \sum_{j=1}^{m} \left[S_1(y_j) - S_2(y_j) \right]^2 \right.$$

$$= \frac{(15)(9)}{(24)^2} \left[0{\cdot}459 + 0.657 \right]$$

$= 0{\cdot}262$ ஆகும்.

$\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில்

$w_{1-\alpha} = w_{\cdot 95} = 0{\cdot}461$

ஆகும்.

T_2-ன் மதிப்பான $0{\cdot}262$, $w_{\cdot 95} = 0{\cdot}461$ஐ விடக் குறைவாக உள்ளதால் சூனிய எடுகோளை ஏற்றுக்கொள்கிறோம்.

மாதிரி: பொது அறிவுத் திறனை அறியும்பொருட்டு ஒரு நகரத்தில் இருந்து 5 மாணவர்களையும், மற்றொரு நகரத்தில் இருந்து 8 மாணவர்களையும் ராண்டம் முறையில் தேர்ந்தெடுத்து

அறிவுத் திறனுக்கான மதிப்பெண்கள் கொடுக்கப்பட்டுக் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டது.

முதற் குழு	இரண்டாம் குழு
82	88
74	77
87	91
86	88
75	94
	93
	83
	94

பொது அறிவுத் திறனில் ஏதாவது வேற்றுமை உள்ளதா என்பதை 'க்ராமர்-வான் மைஸஸ்' சோதனை செய்து அறிக.

X_i	Y_t	$S_1(x) - S_2(x)$
74		$\frac{1}{5} - \frac{0}{8} = \frac{1}{5}$
75		$\frac{2}{5} - \frac{0}{8} = \frac{2}{5}$
	77	$\frac{2}{5} - \frac{1}{8} = \frac{11}{40}$
82		$\frac{3}{5} - \frac{1}{8} = \frac{19}{40}$
	83	$\frac{3}{5} - \frac{2}{8} = \frac{14}{40}$
86		$\frac{4}{5} - \frac{2}{8} = \frac{22}{40}$
87		$\frac{5}{5} - \frac{2}{8} = \frac{30}{40}$
	88	$1 - \frac{3}{8} = \frac{5}{8}$
	88	$1 - \frac{4}{8} = \frac{4}{8}$

X_i	Y_i	$S_1(x) - S_2(x)$
	91	$1 - \frac{5}{8} = \frac{3}{8}$
	93	$1 - \frac{6}{8} = \frac{2}{8}$
	94	$1 - \frac{7}{8} = \frac{1}{8}$
	94	$1 - \frac{8}{8} = 0$

இங்கு $m = 8,\ n = 5$

$$T_2 = \frac{m\,n}{(m+n)^2}\left\{\sum_{i=1}^{n}\Big[S_1(x_i) - S_2(x_i)\Big]^2 + \sum_{j=1}^{m}\Big[S_1(y_j) - S_2(y_j)\Big]^2\right\}$$

$$= \frac{(8)(5)}{(13)^2}\left\{\sum_{i=1}^{n}\Big[S_1(x_i) - S_2(x_i)\Big]^2 + \sum_{j=1}^{m}\Big[S_1(y_j) - S_2(y_j)\Big]^2\right\}$$

இப்பொழுது, $\sum_{i=1}^{5}\Big[S_1(x_i) - S_2(x_i)\Big]^2$

$$= \left\{\left(\frac{1}{5} - \frac{0}{8}\right)^2 + \left(\frac{2}{5} - \frac{0}{8}\right)^2 + \left(\frac{3}{5} - \frac{1}{8}\right)^2 + \left(\frac{4}{5} - \frac{2}{8}\right)^2 + \left(1 - \frac{2}{8}\right)^2\right\}$$

$$= \left(\frac{1}{5}\right)^2 + \left(\frac{2}{5}\right)^2 + \left(\frac{24-5}{40}\right)^2 + \left(\frac{32-10}{40}\right)^2 + \left(\frac{40-10}{40}\right)^2$$

$$= \frac{1}{25} + \frac{4}{25} + \frac{361}{1600} + \frac{484}{1600} + \frac{900}{1600}$$

$$= \frac{5}{25} + \frac{1745}{1600}$$

$$= \frac{320 + 1745}{1600} = \frac{2065}{1600} = 1{\cdot}2906250$$

பிறகு, $\sum_{j=l}^{8} \left[S_1 (y_j) - S_2 (y_j) \right]^2$

$$= \left\{ \left(\frac{2}{5} - \frac{1}{8}\right)^2 + \left(1 - \frac{3}{8}\right)^2 + \left(1 - \frac{4}{8}\right)^2 + \left(1 - \frac{5}{8}\right)^2 + \left(1 - \frac{6}{8}\right)^2 + \left(1 - \frac{7}{8}\right)^2 + \left(1 - \frac{8}{8}\right)^2 \right\}$$

$$= \left\{ \left(\frac{16-5}{40}\right)^2 + \left(\frac{8-3}{8}\right)^2 + \left(\frac{8-4}{8}\right)^2 + \left(\frac{8-5}{8}\right)^2 + \left(\frac{8-6}{8}\right)^2 + \left(\frac{8-7}{8}\right)^2 + \left(\frac{8-8}{8}\right)^2 \right\}$$

$$= \frac{121}{1600} + \frac{25}{64} + \frac{16}{64} + \frac{9}{64} + \frac{4}{64} + \frac{1}{64}$$

$$= \frac{121}{1600} + \frac{55}{64}$$

$$= \frac{1496}{1600} = 0{\cdot}9350000$$

$$T_2 = \frac{(8)\,(5)}{(13)^2} \left\{ 1{\cdot}2906250 + 0{\cdot}9350000 \right\}$$

$$= \frac{40}{169} \left\{ 1{\cdot}2906250 + 0{\cdot}9350000 \right\}$$

$$= (0{\cdot}23668639)\,(2{\cdot}2256250)$$

$$\simeq 0{\cdot}529$$

$\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தைக் கருதினால் $w_{1-\alpha} = w_{0\cdot95}$ன் மதிப்பு $0{\cdot}461$ ஆகும். கண்டறிந்த T_2-ன் மதிப்பு $= 0{\cdot}529$ ஆனது, $w_{\cdot95}$ன் மதிப்பான $0{\cdot}461$ஐ விட அதிகமாக இருப்பதால் H_0ஐ மறுக்கிறோம்.

'வால்டு-வுல்ஃபாவிட்ஸி'ன் ஓட்டச் சோதனை (The Wald-Wolfowitz Runs Test)

இரு சார்பற்ற மாதிரிகள் ஒரே முழுமைத் தொகுதியில் இருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளையும், இரு மாதிரிகளும் மையப் போக்குகளுக்கு அல்லது மாறுபாடுகளுக்கு வேறுபட்டுள்ளன என்ற மாற்று எடுகோளுக்குச் சோதனை செய்ய இச் சோதனை பயன்படுகிறது. பெரும்பாலான சோதனைகள் ஒரு குறிப்பிட்ட வேறுபாட்டை அறியவே பயன்படுகின்றன. உதாரணமாக, இடைநிலைக்குள் வேறுபாடுள்ளனவா என்பதை அறிய இடைநிலைச் சோதனையைத்தான் பயன்படுத்துகிறோம். ஆனால், வால்டு-வுல்ஃபாவிட்ஸ் சோதனையில் எம்மாதிரியான வேறுபாடுகளைச் சோதனை செய்யவும் பயன்படுத்தலாம்.

கருத்தில் கொள்ளப்படும் மாதிரியானது, ஒரு குறிப்பிட்ட தொடர்ச்சியான பரவலைப் பெற்றிருக்க வேண்டும் என்ற எடுகோளையும் இந்த மாறியின் அளவானது குறைந்தது வரிசை அளவாக (ordinal scale) இருத்தல் வேண்டும் என்ற தற்கோளையும் இச் சோதனை கொள்கிறது.

இரு சார்பற்ற மாதிரிகளின் அளவுகள் முறையே n_1, n_2 என்க. இப்பொழுது இவ்விரு குழுக்களின் $n_1 + n_2$ மதிப்புகளை இணைத்து ஏறு வரிசையில் தரப்படுத்துகிறோம். பிறகு இந்த வரிசைப்படுத்தப்பட்ட தொடருக்கான ஓட்டங்களைக் கணக்கிடுகிறோம். உதாரணமாக X, Y என்ற இரு சார்பற்ற குழுக்களின் மதிப்புகள் கீழே கொடுக்கப்பட்டுள்ளன:

X	13	14	7	
Y	12	5	5	2

இங்கு $n_1 = 3, n_2 = 4$. இவ்விரு குழுக்களையும் இணைத்து ஏறுவரிசையில் தரப்படுத்தினால்,

$$\underline{\begin{matrix}2 & 5 & 5\end{matrix}}\quad \underline{7}\quad \underline{12}\quad \underline{\begin{matrix}13 & 14\end{matrix}}$$
$$\begin{matrix}Y & Y & Y & X & Y & X & X\end{matrix}$$

என்று கிடைக்கிறது. இதில் $Y\,Y\,Y$ என்பது ஒரு ஓட்டமாகும். X என்பது இரண்டாவது ஓட்டமாகும். Y என்பது மூன்றாவது ஓட்டமாகும். $X\,X$ என்பது நான்காம் ஓட்டமாகும். இவ்வாறாக மொத்தம் நான்கு ஓட்டங்களைக் காண்கின்றோம். இரண்டு கூறுகளும் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை எனில் X, Y என்ற இருமாதிரிகளும் ஒன்றறக் கலந்திருத்தல் வேண்டும்.

அதாவது, H_0 உண்மையானுல், ஓட்டங்களின் எண்ணிக்கை ஓரளவு பெரியதாக இருக்கும். H_0 தவருனது எனில் ஓட்டங்களின் எண்ணிக்கை சிறியதாக இருக்கும். உதாரணமாக, இரு கூறுகளும் வெவ்வேறு இடைநிலையைப் பெற்ற முழுமைத் தொகுதிகளிலிருந்து எடுக்கப்பட்டவை எனில், ஓட்டங்களின் எண்ணிக்கை சிறியதாக இருக்கும். ஒரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட X என்ற மாதிரியின் இடைநிலையானது மற்ருெரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட Y என்ற மாதிரியின் இடைநிலையைவிட அதிகமாக இருக்கிறது என்று கொள்வோம். இப்பொழுது இரு கூறுகளையும் இணைத்து வரிசைப் படுத்தப்பட்ட மதிப்புகளில் Y-ன் ஓட்டங்கள் இத் தொடரின் இறுதியில் நீண்டும், X-ன் ஓட்டங்கள் அத் தொடரின் ஆரம்பத்தில் நீண்டும் இருக்கும். ஆதலால், ஓட்டங்களின் எண்ணிக்கை ஓரளவு சிறியதாக இருக்கும். மற்ரோர் உதாரணமாக, முழுமைத் தொகுதிகளிலிருந்து தேர்ந்தெடுக்கப்படும் கூறுகளின் மாறுபாடுகளுக்குள் வேறுபாடு இருக்கிறது என்று கொள்வோம். ஒரு முழுமைத் தொகுதியில் இருந்து எடுக்கப்பட்ட X என்ற மாதிரியின் மதிப்புகள் வெகுவாகச் சிதறியும் (dispersed), மற்ருெரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட Y என்ற மாதிரியின் மதிப்புகள் ஒருபடித் தானதாகவும் (homogeneous) இருக்கும். இதனுல், இரு கூறுகளையும் இணைத்து வரிசைப்படுத்தப்பட்ட X-ன் ஓட்டம் நீளமாக இருக்கும். எனவே, ஓட்டங்களின் எண்ணிக்கை ஓரளவு சிறியதாக இருக்கும். பொதுவாக, $r =$ ஓட்டங்களின் எண்ணிக்கை சிறியதாக இருந்தால் H_0-ஐ மறுக்கிருேம். n_1, n_2 அளவுகளைப் பெற்ற வெவ்வேறு கூறுகள் ஒரே தொடராக வரிசைப்படுத்தினுல் r-ன் மாதிரிப் பரவல் உண்டாகிறது. வெவ்வேறு விதமாக வரிசைப் படுத்தினுல், எண்ணிக்கையானது,

$$\binom{n_1 + n_2}{n_1} = \binom{n_1 + n_2}{n_2} \text{ என்று அமையும்.}$$

இதனுல் கண்டறிந்த r-ன் மதிப்புக்குச் சமமாகவோ அல்லது அதற்குக் குறைந்தோ இருப்பதற்கான நிகழ்தகவு,

$$p(r \leqslant r') = \frac{1}{\binom{n_1+n_2}{n_1}} \sum_{r=2}^{r_1} 2 \binom{n_1 - 1}{\frac{r}{2} - 1} \binom{n_2 - 1}{\frac{r}{2} - 1}$$

என்றிருக்கும். இதில் r இரட்டைப்படை எண்ணுகும். r ஒற்றைப்படை எண்ணுகை இருந்தால், நிகழ்தகவு

$$p(r \leqslant r') = \frac{1}{\binom{n_1+n_2}{n_1}} \sum_{r=2}^{n'} \left[\binom{n_1-1}{k-1}\binom{n_2-1}{k-2} + \binom{n_1-1}{k-2}\binom{n_2-1}{k-1} \right]$$

என்றிருக்கும். இதில் $r = 2k - 1$ ஆகும். மேலுள்ள சூத்திரத்தைப் பயன்படுத்தி $n_1, n_2 \leqslant 20$-க்கு r-ன் தீர்வு கட்டமான மதிப்புள்ள அட்டவணை F_1-ல் கொடுக்கப்பட்டுள்ளது. இம் மதிப்புகள் $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்திற்கு ஏற்ற மதிப்புகளாகும். கண்டறிந்த r-ன் மதிப்பு, n_1, n_2-க்கு ஏற்ற அட்டவணையின் மதிப்புக்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால், H_0-ஐ $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். கண்டறிந்த r-ன் மதிப்பு அட்டவணையின் மதிப்புக்கு அதிகமாக இருந்தால், $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் சூனிய எடுகோளை மறுப்பதற்கில்லை என்று தீர்மானிக்கிறோம்.

மாதிரி : ஆறு வயதுடைய 12 சிறுவர்களும், 12 சிறுமிகளும் அவர்கள் சேர்ந்து விளையாடும் போக்கில் ஆக்கிரமிப்புத் தன்மையானது இருபாலாரிடத்தும் ஒரே மாதிரியாக இருக்கிறதா அல்லது ஆக்கிரமிப்புத் தன்மையில் இன வேறுபாடு உள்ளதா என்பதை அறிய அவர்களின் ஆக்கிரமிப்பு மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டன.

சிறுவர்கள் (X)	சிறுமியர்கள் (Y)
97	66
80	51
83	33
76	69
124	27
76	18
129	20
56	27
152	37
115	47
52	31
61	26

இப்பொழுது 'வால்டு-வுல்ஃபாவிட்ஸ்' சோதனையைப் பயன்படுத்தி இரு பாலாரிடத்தும் ஆக்கிரமிப்புத்தன்மை ஒரே மாதிரியாக உள்ளது என்ற சூனிய எடுகோளையும், ஆக்கிரமிப்புத் தன்மையில் இருபாலாரும் வேறுபடுகின்றனர் என்ற மாற்று எடுகோளையும் கருதி ஆய்வு செய்வோம். $\alpha = 0{\cdot}05$ என்ற மிகைத்

தன்மை மட்டத்தைக் கருதுக. இங்கு $n_1 = 12, n_2 = 12$ ஆகும். கொடுக்கப்பட்ட இரு குழுவையும் இணைத்து வரிசைப்படுத்தினால்,

18	20	26	27	27	31	33	33	47	51
X	X	X	X	X	X	X	X	X	X

(1)

52	56	61
Y	Y	Y

(2)

66	69
X	X

(3)

76	76	80	83	97	115	124	129	152
Y	Y	Y	Y	Y	Y	Y	Y	Y

(4)

இங்கு மொத்த ஓட்டங்கள் $r = 4$ ஆகும். அட்டவணை F_1-லிருந்து $n_1 = 12, n_2 = 12$-க்கான r-ன் தீர்வு கட்டமான மதிப்பு $= 7$ ஆகும்.

கண்டறிந்த r-ன் மதிப்பான 4, அட்டவணை மதிப்பான 7ஐ விடக் குறைவாக இருப்பதால், H_0 ஐ $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

கவனிக்கவும்: n_1 அல்லது $n_2 > 20$ எனில், அட்டவணை F_1 ஐப் பயன்படுத்த முடியாது, எனினும் இம்மாதிரியான பெரிய கூறு களுக்கு H_0-ன் கீழ் r-க்கான மாதிரிப் பரவலானது தோராயமாக, இயல்நிலையாக,

$$\text{சராசரி} = \mu = \frac{2\, n_1\, n_2}{n_1 + n_2} + 1$$

$$\text{திட்டவிலக்கம்} = \sigma = \sqrt{\frac{2\, n_1\, n_2\,(2\, n_1\, n_2 - n_1 - n_2)}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)}}$$

என்ற சுட்டுறுப்புகளைப் பெற்று அமைகிறது.

அதாவது, $z = \dfrac{r - \mu}{\sigma}$

$$= \frac{r - \left[\dfrac{2\, n_1\, n_2}{n_1 + n_2} + 1\right]}{\sqrt{\dfrac{2\, n_1\, n_2\,(2\, n_1\, n_2 - n_1 - n_2)}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)}}}$$

ஆனது பூச்சிய சராசரியையும் ஒருமை மாறுபாட்டையும் பெற்று இயல்நிலைப் பரவலாக அமைகிறது. அட்டவணை A-ல் z-க்கு தொடர்புடைய H_0-ன் கீழ் நிகழ்ச்சி நிகழ்வதற்கான நிகழ்தகவு

மதிப்புகள் கொடுக்கப்பட்டுள்ளன. (n_1+n_2)-ன் மதிப்பு அவ்வளவு பெரியதாக இல்லையெனில் தொடர்ச்சிக்கான திருத்தம் செய்தல் அவசியமாகிறது. இவ்வாறாக z ஆனது,

$$z = \frac{|r-\mu| - 0\cdot5}{\sigma}$$ என்று துல்லியமாக்கப்படுகிறது.

அதாவது, $$z = \frac{\left| r - \left(\frac{2\,n_1\,n_2}{n_1+n_2} + 1\right)\right| - 0\cdot5}{\sqrt{\frac{2n_1\,n_2\ \ (2\,n_1\,n_2 - n_1 - n_2)}{(n_1+n_2)^2\,(n_1+n_2-1)}}}$$

மேலுள்ள சூத்திரத்தின் மூலம் பெறப்பட்ட z-ன் மதிப்புக்குத் தொடர்புடைய நிகழ்தகவின் மதிப்பான p-யானது, α என்ற மிகைத்தன்மை மட்டத்திற்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால் H_0-ஐ α மிகைத்தன்மை மட்டத்தில் நாம் மறுக்கிறோம்.

மாதிரி: 21 சாதாரண நிலையிலுள்ள கற்றுக் கொடுக்கப்பட்ட முயல்களும், மூளை ஆபரேஷன் செய்யப்பட்ட 8 கற்றுக் கொடுக்கப்பட்ட முயல்களும் அவைகளின் ஞாபக சக்தியின் மதிப்புகளைக் கீழ்க்கண்டவாறு பெற்றுள்ளன.

ஆபரேஷன் செய்யப்பட்ட முயல்கள்	சாதாரண நிலையில் உள்ள முயல்கள்
30	33
65	18
39	34
34	25
85	18
66	16
41	25
55	25
	31
	33
	26
	25
	34
	25
	31
	25
	28
	24
	32
	25
	24

சாதாரண நிலையிலுள்ள முயல்களுக்கும், ஆபரேஷன் செய்யப் பட்ட முயல்களுக்குமிடையே வேறுபாடு இல்லை என்ற சூனிய எடு கோளையும், இரு குழுக்களுக்கிடையே வேறுபாடு உள்ளது என்ற மாற்று எடுகோளையும் கருதி வால்டு-வுல்ஃபாவிட்ஸ் சோதனை மூலம் ஆய்வு செய்வோம். $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். கொடுக்கப்பட்ட இரு கூறுகளையும் இணைத்து வரிசைப்படுத்தி ஓட்டங்களின் எண்ணிக்கையைக் காண்போம். அவ்வாறே,

16	18	18	24	24	25	25	25	25	25	25	25	26	28
Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

(1)

30	31	31	32	33	33	34	34	34	39	41	55	65	66	85
X	Y	Y	Y	Y	Y	Y	X	Y	X	X	X	X	X	X
(2)	(3)						(4)	(5)	(6)					

ஓட்டங்களின் மொத்த எண்ணிக்கை $r = 6$ ஆகும். இங்கு $n_1 = 8$, $n_2 = 21$ ஆகும்.

$n_2 > 20$ என்பதாலும், $n_1 + n_2 = 29$ ஆனது ஓரளவு சிறியது என்பதாலும் தொடர்ச்சியான திருத்தத்தைப் பெற்ற சூத்திர மான,

$$z = \frac{\left| r - \left(\dfrac{2\,n_1\,n_2}{n_1 + n_2} + 1 \right) \right| - 0{\cdot}5}{\sqrt{\dfrac{2\,n_1\,n_2\,(2\,n_1\,n_2 - n_1 - n_2)}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)}}}$$

$$= \frac{\left| 6 - \left(\dfrac{(2)(8)(21)}{8 + 21} + 1 \right) \right| - 0{\cdot}5}{\sqrt{\dfrac{(2)(8)(21)\{(2)(8)(21) - 8 - 21}{(8 + 21)^2\,(8 + 21 - 1)}}}$$

$= 2{\cdot}92$ ஆகும்.

அட்டவணை A-லிருந்து H_0-ன் கீழ் $z \geqslant 2{\cdot}92$-க்கு ஆன நிகழ்தகவு $p = 0{\cdot}0018$ ஆகும்.

இது $\alpha = 0.01$ ஐ விடக் குறைவாகவே இருப்பதனுல் H_0-ஐ மறுக்கிறோம்.

குறிப்பு : சாதாரணமாக ஓட்டச் சோதனையில் ஒத்தவை (ties) நிகழ்வது மிக அபூர்வமாகும். சில சமயங்களில் ஒத்தவை நிகழ்வதுண்டு. உதாரணமாக, கூறு X-ல் உள்ள மதிப்பு 16 16 என்றும், கூறு Yலுள்ள மதிப்பு 16 என்றும் இருப்பின் இதனை வரிசைப்படுத்தும்போது XXX என்றும், XXY என்றும், YXX என்றும் வரிசைப்படுத்தலாம். இதனால் ஓட்டங்களின் எண்ணிக்கையில் மாற்றமிருக்கும். மேலே கொடுக்கப்பட்டுள்ள மாதிரியில் இவ்வாறான ஒத்தவை நிகழ்ந்துள்ளன. அதாவது,

$$\begin{matrix} 34 & 34 & 34 \\ Y & X & Y \end{matrix}$$

என்று அமைந்துள்ளது. இதனை $X\,Y\,Y$ என்றும் வரிசைப்படுத்தலாம். ஆனால், எப்படி வரிசைப்படுத்தினாலும் r-ன் எண்ணிக்கை 6 ஆகவோ அல்லது அதற்குக் குறைவாகவோ இருக்கும் என்று அறியப்படுகிறது. எப்படியும் இதனால் H_0 மறுக்கப்படுகிறது என்பதால் இந்த ஒத்தவைகளைச் சாதாரணமாகக் கருதி ஓட்டங்களின் எண்ணிக்கையைக் கணக்கிட்டோம். சில சமயங்களில், மேற்கூறியபடி சாதகமாக அமையாமல் தீர்மான விதியை அல்லது உய்த்துணர்வை பாதிக்கலாம். அச் சமயத்தில், $Y\,X\,Y$ என்று கொண்டு ஓட்டங்களின் மொத்த எண்ணிக்கையை அறிந்து z_1-ஐக் கண்டபின் நிகழ்தகவு p_1-ன் மதிப்பை அட்டவணை A-விருந்து காணவேண்டும். பிறகு $X\,Y\,Y$ என்று கொண்ட ஓட்டங்களின் மொத்த எண்ணிக்கையை அறிந்து, z_2-ஐக் கண்ட பிறகு, நிகழ்தகவு p_2-ன் மதிப்பைக் காணவேண்டும். பிறகு $Y\,Y\,X$ என்று கொண்டு ஓட்டங்களின் மொத்த எண்ணிக்கையை அறிந்து z_3-ஐக் கண்டபின் நிகழ்தகவு p_3-ன் மதிப்பைக் காணவேண்டும். முடிவில், p_1, p_2, p_3 ஆகியவைகளின் சராசரியின் மதிப்பானது α என்ற மிகைத்தன்மை மட்டத்திற்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால், H_0-ஐ மறுக்கிறோம்.

'மோஸஸி'ன் சோதனை (Moses Test)

X என்பது கட்டுப்பட்ட வகையைச் (Control Cases) சேர்ந்த குழு என்றும் Y என்பது ஆராய்ச்சி வகையைச் சேர்ந்த குழு என்றும் கொள்வோம். n_X கட்டுப்பட்ட வகைகளும், n_Y ஆராய்ச்சி வகைகளும் கொண்ட குழுக்களைக் கருதி $n_X + n_Y$-ன் மதிப்புகளை ஏறு வரிசையில் அமைக்கிறோம். Xகளும், Yகளும், ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோள் உண்மையெனில் ஒரே வரிசையில் அமைக்கப்பட்ட Xகளும் Yகளும் ஒன்றறக் கலந்திருக்கும் என்று எதிர்பார்க்கிறோம். சூனிய எடுகோள் H_0-ன் கீழ் சில மிக அத்துமீறிய

மதிப்புகள் (extremely high scores) Xகளாகவும், சில Yகளாகவும் அல்லது சில மிகச் சிறிய மதிப்புகள் Xகளாகவும், சில Yகளாக வும், மத்திய வீச்சிலுள்ள மதிப்புகள் X, Y ஆகியவைகளின் கலவை யாகவும் இருக்கும் என்று எதிர்பார்க்கிறோம். மாற்று எடுகோள் H_1 உண்மையானால் பெரும்பாலான (i) X-களின் மதிப்புகள் குறைந்திருக்கும் என்றோ அல்லது (ii) பெரும்பாலான X-களின் மதிப்புகள் அதிகமாக இருக்கும் என்றோ அல்லது (iii) X-களின் ஓரளவு மதிப்புகள் குறைந்தும் அதே சமயத்தில் X-களின் மற்ற மதிப்புகள் அதிகமாகவும் இருக்கும் என்று எதிர்பார்க்கிறோம். மாற்று எடுகோள் H_1 உண்மையானால் பெரும்பாலான (i) X-களின் மதிப்புகள் குறைந்திருக்கும் என்றோ அல்லது (ii) பெரும் பாலான X-களின் மதிப்புகள் அதிகமாக இருக்கும் என்றோ அல்லது (iii) X-களின் ஓரளவு மதிப்புகள் குறைந்தும் அதே சமயத்தில் X-களின் மற்ற மதிப்புகள் அதிகமாகவும் இருக்கும் என்று எதிர் பார்க்கிறோம்.

மாற்று எடுகோள் (i) உண்மையானால் X, Y ஆகியவைகளை இணைத்து வரிசைப்படுத்தப்பட்ட தொடரின் இறுதிப் பகுதியில் X-களின் மதிப்புகள் மிக நெரிசலாகக் காணப்படும். மாற்று எடுகோள் (ii) உண்மையானால் X, Y ஆகியவைகளை இணைத்து வரிசைப்படுத்தப்பட்ட தொடரின் ஆரம்பத்தில் X-களின் மதிப்புகள் மிக நெரிசலாகக் காணப்படும். மாற்று எடுகோள் (iii) உண்மையானால் X, Y ஆகியவைகளை இணைத்து வரிசைப் படுத்தப்பட்ட தொடரின் மத்தியில் X-களின் மதிப்புகள் நெரிச லோடு காணப்படும்.

X-களின் மதிப்புகள் $n_X + n_Y$ மதிப்புகளுக்கு மிக நெருங்கிய முறையில் நெரிசலாக உள்ளதா என்பதை அறிய மோஸஸ் சோதனை பயன்படுகிறது.

X, Y ஆகிய இரு குழுக்களின் மதிப்புகளையும் இணைத்து வரிசைப்படுத்தவேண்டும். வரிசைப்படுத்தப்பட்ட தொடரின் மதிப்புகள் குழுவின் அடையாளத்தைக் குறிக்கும். இணைக்கப் பட்டு வரிசைப்படுத்தப்பட்ட தொடரில் X-ன் பரவுதலின் மதிப் பைக் கணக்கிடுகின்றோம். இதனை s' என்று குறிப்பிடுகிறோம். X-ன் மிகக் குறைந்த மதிப்பையும், X-ன் மிக உயர்ந்த மதிப்பையும் அறிந்து அவைகளுக்கிடையேயுள்ள தரங்களின் எண்ணிக்கையே X-ன் பரவுதலின் மதிப்பாகும்.

உதாரணமாக $n_X = 6, n_Y = 7$ என்க.

இவ்விரு குழுக்களையும் இணைத்து வரிசைப்படுத்தியதில் கீழ்க்கண்டவாறு அமைகிறது என்க.

தரம் (Rank):	1	2	3	4	5	6	7	8	9	10	11	12	13
குழு :	Y	Y	X	Y	X	Y	X	X	X	Y	X	Y	Y

இதில் X மதிப்புகளின் பரவுதலான s' ஆனது தரம் 3லிருந்து தரம் 11 வரை பரவியுள்ளது. தரம் 3லிருந்து தரம் 11 வரையுள்ள மொத்த எண்ணிக்கைகள் 9 ஆகும். எனவே $s' = 9$ ஆகும்.

இதனை எளிதாகக் கீழ்க்கண்டவாறும் அறியலாம். X மதிப்புகளின் பரவுதல் தரம் 3லிருந்து தரம் 11 வரை உள்ளது.

$s' = X$ பரவுதலின் வீச்சு $+ 1$ ஆகும்.

$= (11 - 3) + 1 = 9$ ஆகும்.

X-களும், Y-களும் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்பதைத் தெரிவிக்கும் வகையில் கண்டறிந்த s'-ன் மதிப்பு மிகச் சிறியதாக உள்ளதா என்பதை இச் சோதனை அறிவிக்கிறது. அதாவது H_0-ன் கீழ் s'-ன் மாதிரிப் பரவல் தெரிந்துள்ளது என்றும் அதனால் இதனைப் பயன்படுத்தி மிகைத்தன்மைச் சோதனை செய்யலாம்.

n_X-ன் அளவு மிகப் பெரியதாக இருந்தால் இச் சோதனைக்குத் திருத்தப்பட்ட மாறுதல் அவசியம் என்றும் மோஸஸ் கண்டறிந்துள்ளார்.

விவரங்களைக் கண்டறியும் முன்பே h என்ற ஒரு சிறிய எண்ணை யாதாமொரு முறையில் கொள்ளவேண்டும். பின்பு X-ன் பரவுதல், ஏற்படும் தரத்திலிருந்து முடிவடையும் தரம் வரை அறிந்த பின்னர், h-க்கு ஏற்றவாறு அப் பரவுதலின் இரு நுனிகளிலும் உள்ள தரங்களை நீக்கிவிட வேண்டும்.

மேலே சொன்ன உதாரணத்தில், X-ன் பரவுதல் தரம் 3லிருந்து தரம் 11 வரை உள்ளது. இப்பொழுது $h = 1$ எனக் கொண்டால் X பரவுதலின் நுனித் தரங்களான 3-ஐயும், 11-ஐயும் நீக்கிவிடவேண்டும். பின்பு X-ன் பரவுதலான s_h காண விழைகிறோம். இப்பொழுது X-ன் பரவுதல் தரம் 5லிருந்து தரம் 9 வரை நிகழ்கிறது. எனவே. $s_h = 9 - 5 + 1 = 5$.

$h = 2$ எனக் கொண்டிருந்தால் $\overset{3}{X}$ $\overset{5}{X}$ என்ற இடது நுனித் தரங்களையும் $\overset{9}{X}$ $\overset{11}{X}$ என்ற வலது நுனித்தரங்களையும் நீக்கிவிடு கிறோம். பிறகு X-ன் பரவுதல் தரம் 7லிருந்து தரம் 8 வரை நிகழ்கிறது. எனவே, $s_h = 8 - 7 + 1 = 2$ ஆகும்.

குறிப்பு: s_h ஆனது $(n_X - 2h)$-ஐவிடக் குறைவாக இருக்காது. அதே சமயத்தில் s_h ஆனது $(n_X + n_Y - 2h)$ ஐ விட அதிகமாகவும் இருக்காது.

H_0-ன் கீழ் s_h-ன் கண்டறிந்த மதிப்பு $(n_X - 2h)$-ன் நீச மதிப்பை விட அதிகமாக இருப்பதற்கான நிகழ்தகவு மாதிரிப் பரவல் நமக்குத் தெரிவிக்க வேண்டும். s_h ஆனது $(n_X - 2h)$-ன் நீச மதிப்பைவிட அதிகமாக இருக்கும் மதிப்பு g என்று குறிப்போம்.

இப்பொழுது H_0-ன் கீழ் s_h ஆனது $(n_X - 2h + g)$ மதிப்புக்குச் சமமாக அல்லது குறைவாக இருப்பதற்கான நிகழ்தகவு,

$$p\{s_h \leqslant n_X - 2_h + g\}$$

$$= \left\{ \frac{\sum_{k=0}^{g} \binom{k + n_X - 2h + g}{k} \binom{n_Y + 2h + 1 - k}{n_Y - k}}{\binom{n_X + n_Y}{n_X}} \right\}$$

p-ன் மதிப்பானது $\alpha = 0{\cdot}05$-க்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால் X-களும், Y-களும் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளை $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

p-ன் மதிப்பானது $\alpha = 0{\cdot}05$ஐவிட அதிகமாக இருந்தால் கொடுக்கப்பட்ட விவரமானது H_0-ஐ மறுப்பதற்கு அனுமதிக்க வில்லை என்று $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் உய்த்துணரு கிறோம்.

மாதிரி: திரைப்படங்களில் வரும் சண்டைக் காட்சிகளை எவ்விதமான கலவரமடையாத நிலையிலோ அல்லது கலவர மடைந்த நிலையிலோ பார்க்கும் இரு குழுக்களின் ஆளுமைச் சோதனை (personality test) மதிப்புகளை ஒத்துப்பார்த்ததில்

X என்ற குழுவிலுள்ளவர்கள் கலவரமடையாத நிலையிலும், அதாவது கட்டுப்பட்ட நிலையிலும் (controlled state) *Y* என்ற குழுவிலுள்ளவர்கள் கலவரமடைந்த நிலையிலும் இருந்ததாக அறியப்பட்டது. *X* குழுவிலுள்ள 9 பேருக்கும், *Y* குழுவிலுள்ள 9 பேருக்கும் சண்டைக் காட்சிகள் காட்டப்பட்டு அவர்களின் ஆளுமை மதிப்புகளை அவர்களே குறித்துக்கொள்ளும்படி கேட்டுக்கொள்ளப்பட்டனர். அம்மாதிரி மதிப்புகள் கீழ்க்கண்ட அட்டவணையில் உள்ளன.

X	*Y*
12	25
16	5
6	14
13	19
13	0
3	17
10	15
10	8
11	8

இப்பொழுது *X* குழுவிலுள்ள கலவரமடையாதவர்களின் ஆளுமைத் திறனுக்கும், *Y* குழுவில் உள்ள கலவரமடைபவர்களின் ஆளுமைத்திறனுக்கும் வேறுபாடு இல்லை என்ற சூனிய எடுகோளையும் *X* குழுவிலுள்ள கலவரமடையாதவர்களின் ஆளுமைத் திறன், *Y* குழுவிலுள்ள கலவரமடைபவர்களின் ஆளுமைத் திறனை விட அதிகமாக உள்ளது என்ற மாற்று எடுகோளையும் கொண்டு 'மோஸஸ்' சோதனையைச் செய்வோம்.

$\alpha = 0{\cdot}05$ என்க; $n_X = 9$, $n_Y = 9$ ஆகும். இரு குழுக்களையும் இணைத்து வரிசைப்படுத்தி, தரங்களைக் கொடுத்து அக் குழுக்களின் அடையாளங்களைக் குறித்தால்,

மதிப்பு :	0,	3,	5,	6,	8,	8,	10,	10,	11,	12,	13,	13,	14,	15,	16,	17,	19,	25
தரம் :	1,	2,	3,	4,	5,	6,	7,	8,	9,	10,	11,	12,	13,	14,	15,	16,	17,	18
குழு :	*Y*,	*X*,	*Y*,	*X*,	*Y*,	*Y*,	*X*,	*X*,	*X*,	*X*,	*X*,	*X*,	*Y*,	*Y*,	*X*,	*Y*,	*Y*,	*Y*

என்று பெறுகிறோம்.

இங்கு $h = 1$ என்று கொண்டால், *X*-ன் பரவுதலின் தரம் 2லிருந்து, தரம் 15 வரை நிகழ்கிறது. நுனித் தரங்களான 2-ஐயும் 15-ஐயும் நீக்கியபின்னர் *X*-ன் பரவுதலின் தரம், 4-லிருந்து,

தரம் 12 வரை நிகழ்கிறது. எனவே, $s_h = 12 - 4 + 1 = 9$ ஆகும். $h = 1$ என்பதால், s_h-ன் முடியுமான மிகக் குறைந்த மதிப்பு $n_X - 2h = 9 - (2)(1) = 7$ ஆகும்.

s_h-ன் கண்டறிந்த மதிப்பு இந்தக் குறைந்த மதிப்பைவிட 2 அதிகமாக உள்ளது. அதாவது,

$$g = s_h - (n_X - 2h)$$
$$= 9 - 7 = 2 \text{ ஆகும்.}$$

இப்பொழுது $n_X = 9, n_Y = 9, g = 2$-க்கு

H_0-ன் கீழ் $s_h \leqslant 9$-க்கான நிகழ்தகவு

$$p\,(s_h \leqslant n_X - 2h + g)$$

$$= \frac{\sum_{k=0}^{g} \binom{k + n_X - 2h - 2}{k} \binom{n_Y + 2h}{n_Y - k}}{\binom{n_X + n_Y}{n_X}}$$

$$= \frac{\sum_{k=0}^{2} \binom{k + 9 - 2 - 2}{k} \binom{9 + 2 + 1 - k}{9 - k}}{\binom{9 + 9}{9}}$$

$$= \frac{\binom{5}{0}\binom{12}{9} + \binom{6}{1}\binom{11}{8} + \binom{7}{2}\binom{10}{7}}{\binom{18}{9}}$$

$$= \frac{1\,(220) + 6\,(165) + 21\,(120)}{48620}$$

$$= 0{\cdot}077$$

அதாவது $p = 0{\cdot}077$ ஆகும். இது $\alpha = 0{\cdot}05$ஐ விட அதிகமாக உள்ளது. எனவே, கொடுக்கப்பட்ட விவரமானது H_0-ஐ மறுப்பதற்கு அனுமதிக்கவில்லை என்று $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் உய்த்துணருகிறோம்.

குறிப்பு : ஒத்தவை இருந்தால் வால்டு - வுல்ஃபாவிட்ஸ் சோதனையில் கண்டதுபோல் முடிந்த வெவ்வேறு s_h-ன் மதிப்புகளைக்

கண்டு வெவ்வேறு p-க்களையும் காணவேண்டும். பிறகு இந்த p-க்களின் சராசரி காணவேண்டும். நிறைய ஒத்தவைகளிருந்தால் மோஸஸ் சோதனையைப் பயன்படுத்த இயலாது.

இரு சார்பற்ற கூறுகளுக்கு ராண்டம் மயமான சோதனை (Randomisation Test for Two Independent Samples)

இரு சார்பற்ற குழுக்களில் முதற் குழுவின் அளவான n_1-ம், இரண்டாம் குழுவின் அளவான n_2-ம் மிகச்சிறியவைகளாக இருந்து இவ்விரு குழுக்களின் சராசரிகளுக்கிடையே வேறுபாடு உள்ளதா இல்லையா என்பதை இச் சோதனை மூலம் அறியலாம். H_0-ன் கீழ் கண்டறிந்தவைகளோடு தொடர்புடைய சரியான நிகழ்தகவு மதிப்புகளை ராண்டம் மயமான சோதனை மூலம் காணலாம். இதில் இச் சோதனை இயல்நிலைப் பரவல் என்ற தற்கோளையோ மாறுபாடுகள் ஓரினமுடையன (Homogenity of Variance) என்ற தற்கோளையோ கொள்வதில்லை.

இரு முழுமைத் தொகுதிகளிலிருந்து ராண்டம் முறையில் எடுக்கப்பட்ட இரு சார்பற்ற கூறுகளைக் கருதுக. உதாரணமாக, முதற்குழுவின் அளவு $n_1 = 4$, இரண்டாம் குழுவின் அளவு $n_2 = 5$ என்று கீழ்க்காணும் விவரங்களைக் காண்போம்.

குழு I :	1	12	13	21	
குழு II :	17	20	23	25	30

இந்தக் கண்டறிந்த மதிப்புகளைக் கொண்டு, சராசரிகளுக்குள் எவ்வித வேறுபாடும் இல்லை என்ற சூனிய எடுகோளைக் குழு I-ன் சராசரி, குழு II-ன் சராசரியை விடக் குறைவானது என்ற மாற்று எடுகோளுக்கு எதிராகச் சோதனை செய்ய விரும்புகிறோம். சூனிய எடுகோளின் கீழ் எல்லா n_1, n_2 கண்டறிந்தவைகளும் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்று கொள்ளுகிறோம்.

$$\binom{n_1 + n_2}{n_1} = \binom{4 + 5}{4} = 126$$

H_0-ன் கீழ் குழு I, குழு II என்று மதிப்புகளுக்குச் சமவாய்ப்பு வழிகளில் ஏதாவதொரு வழியில் நிர்ணயிக்கப்பட்டிருக்கலாம்.

H_0-ன் கீழ் 126 முயற்சிகளில் ஒரு முறை மட்டும் 9 மதிப்புகளில் 4 சிறிய மதிப்புகள் குழு I என்றும், மற்றப் பெரிய மதிப்புகள்

குழு II என்றும் குறிக்கப்படும். ஒரே முயற்சியில் இம்மாதிரியான நிகழ்ச்சி நிகழுமானால் $p = 1/126 = 0{\cdot}008$ என்ற மிகைத் தன்மை மட்டத்தில் H_0-ஐ மறுப்போம். இரு குழுக்களும் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றதெனில் 126 சாத்திய மான வெளிப்பாடுகளில் மிக அத்துமீறியவை (most extreme) மட்டும் நிகழமுடியும் என்று கருத முடியாது.

இந்த ராண்டடம் மயமான சோதனையானது சாத்தியமான அத்துமீறிய $(n_1 + n_2)$ மதிப்புகளோடு வெளிப்பாடுகளின் எண்ணிக்கையைத் தெரிவிக்கிறது. இவ் வெண்ணிக்கைகளை மறுக்கத்தகு வெளிகள் என்று கூறுகிறோம்.

H_0-ன் கீழ் $\binom{n_1 + n_2}{n_1}$ சம வாய்ப்பு நிகழ்ச்சிகள் நிகழ்ந் தால் மொத்த மதிப்புகளுக்கும், இரண்டாம் குழுவின் மொத்த மதிப்புகளுக்குமிடையேயுள்ள வித்தியாசம் அத்துமீறிய தாக (extreme) இருக்கும். முதல் குழுவை X என்றும், இரண்டாம் குழுவை Y என்றும் கொண்டால், ΣX-க்கும், ΣY-க்கும் இடையே யுள்ள வித்தியாசம் அத்துமீறியதாக இருக்கும். இவைகளின் வித்தியாசங்கள் பெரியதாக இருக்கும் வகைகள் மறுக்கத்தகு வெளியைக் குறிக்கும். α என்ற மிகைத்தன்மை மட்டத்தினைக் கருதினால், மறுக்கத்தகு வெளியானது (Rejection Region), நிகழும் நிகழ்ச்சிகளில், $\alpha \binom{n_1 + n_2}{n_1}$ பெரும்பாலான மிக அத்துமீறிய நிகழ்ச்சிகளைக் கொண்டு இருக்கும். அதாவது, $\alpha \binom{n_1 + n_2}{n_1}$ எண்ணிக்கையிலான வெளிப்பாடுகள் மறுக்கத்தகு வெளியைக் குறிக்கும்.

சில குறிப்பிட்ட வெளிப்பாடுகளில் X-களின் சராசரி Y-களின் சராசரியை விட அதிகமாக இருக்கும். இந்த வெளிப்பாடுகள் ΣX-க்கும், ΣY-க்கும் இடையே உள்ள வித்தியாசம் மிக அதிக மாக இருக்கும்போது ஏற்படுகிறது. நாம் கருதும் கூறு மேற் சொன்ன மறுக்கத்தகு வெளியில் அமையுமானால் H_0-ஐ α என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். மேலே கூறிய உதாரணத்தில் 9 மதிப்புகளுள்ளன. மேலே ΣX-க்கும், ΣY-க்கும் இடையேயுள்ள சாத்தியமான வித்தியாசங்கள்

$$\binom{n_1 + n_2}{n_1} = \binom{4 + 5}{4} = 126 \text{ உள்ளன.}$$

$\alpha = 0 \cdot 05$ என்று கருதினால், மறுக்கத்தகு வெளியானது (Rejection Region), $\alpha \binom{n_1 + n_2}{n_1}$

$= 0 \cdot 05\ (126) = 6 \cdot 3.$

ΣX, ΣY ஆகியவைகளின் வித்தியாசங்களின் திசையை எதிர்பார்க்காமல் மாற்று எடுகோளைக் கொண்டிருந்தால் H_0-க்கு இருமுனைச் சோதனையைப் பயன்படுத்துகிறோம். அப்பொழுது 6 அத்துமீறிய சாத்தியமான வெளிப்பாடுகளில் 3 அத்துமீறிய சாத்தியமான வெளிப்பாடுகள் அதற்கு எதிரான திசையிலும் இருக்கும். இது 6 சாத்தியமான வெளிப்பாடுகளுக்கு ΣX, ΣY களுக்கான தனி வித்தியாச சாத்தியமான உச்சத்தைப் பெற்றதாக அமையும், அத்துமீறிய வெளிப்பாடுகளைப் பெற்றிருக்கும். எனவே, மறுக்கத்தகு வெளியானது 6 அத்துமீறிய சாத்தியமான வெளிப்பாடுகளை ஒரு குறிப்பிட்ட திசையில் பெற்றிருக்கும். $\mu_X < \mu_Y$ என்ற மாற்று எடுகோளின் கீழ் $\alpha = 0 \cdot 05$ என்ற மறுக்கத்தகு வெளியை உண்டாக்கும் 6 சாத்தியமான வெளிப்பாடுகள் பின்வருமாறு :

சாத்தியமான 5Y-வகைகளின் மதிப்புகள்	சாத்தியமான 4X-வகைகளின் மதிப்புகள்	$\Sigma Y - \Sigma X$
20, 21, 23, 25, 30	1, 12, 13, 16	119 − 43 = 76
17, 21, 23, 25, 30	1, 12, 13, 20	116 − 46 = 70
17, 20, 23, 25, 30	1, 12, 13, 21	115 − 47 = 68
17, 20, 21, 25, 30	1, 12, 13, 23	113 − 49 = 64
13, 21, 23, 25, 30	1, 12, 17, 20	112 − 50 = 62
17, 20, 21, 23, 30	1, 12, 13, 25	111 − 51 = 60

இதில் $Y = 17, 20, 23, 25, 30$; $X = 1, 12, 13, 21$ என்பன நம்முடைய கூறுகளின் கண்டறிந்த மதிப்புகளாகும். கண்டறிந்தவை மறுக்கத்தகும் வெளியில் இருப்பதால் H_0-ஐ

மறுக்கிறோம். கண்டறிந்த மதிப்புகள் அல்லது H_0-ன் கீழ் கட்டான அத்துமீறியவை நிகழ்வதற்கு, ஒருமுனைக்கான நிகழ்தகவு $p = \frac{3}{126}$ ஆகும்.

ΣX, ΣY ஆகியவைகளின் வித்தியாசங்களின் திசையை எதிர்பார்க்காமல், மாற்று எடுகோளைக் கொண்டிருந்தால், H_0-க்கு இருமுனைச் சோதனையைப் பயன்படுத்துகிறோம். அப்பொழுது, 6 அத்துமீறிய, சாத்தியமான வெளிப்பாடுகளில் 3 அத்துமீறிய சாத்தியமான வெளிப்பாடுகள் ஒரு திசையிலும், மற்ற 3 அத்துமீறிய சாத்தியமான வெளிப்பாடுகள் அதற்கு எதிரான திசையிலும் இருக்கும். 6 சாத்தியமான வெளிப்பாடுகளுக்கு, ΣX, ΣY களுக்கான, தனி வித்தியாசங்கள் (absolute differences) உச்சத்தைப் பெற்றதாக அமையும். உதாரணமாக, 6 அத்துமீறிய, சாத்தியமான வெளிப்பாடுகள், $\alpha = 0{\cdot}05$ என்ற மிகைத் தன்மை மட்டத்தில், இருமுனைச் சோதனைக்குக் கீழ்க்கண்டவாறு அமையும்.

சாத்தியமான 5Y-வகைகளின் மதிப்புகள்	சாத்தியமான 4X-வகைகளின் மதிப்புகள்	$\lvert \Sigma Y - \Sigma X \rvert$
20, 21, 23, 25, 30	1, 12, 13, 17	$\lvert 119 - 43 \rvert = 76$
1, 12, 13, 17, 20	21, 23, 25, 30	$\lvert 63 - 99 \rvert = 36$
17, 21, 23, 25, 30	1, 12, 13, 20	$\lvert 116 - 46 \rvert = 70$
1, 12, 13, 17, 21	20, 23, 25, 30	$\lvert 64 - 98 \rvert = 34$
17, 20, 23, 25, 30	1, 12, 13, 21	$\lvert 115 - 47 \rvert = 68$
1, 12, 13, 17, 23	20, 21, 25, 30	$\lvert 66 - 96 \rvert = 30$

நிகழ்தகவு மதிப்புகள் அட்டவணை B-ல் கொடுக்கப்பட்டுள்ளன. முழுமைத் தொகுதிகளானவை, பொதுவாக மாறுபாட்டைப் பெற்ற இயல்நிலைப் பரவலிலிருந்து வரப்பெற்றவை என்ற தற்கோளைக் (assumption) கொள்ளவில்லை என்ற காரணத்தாலும், இங்கு இச் சோதனையானது, சுட்டுறுப்பைச் சார்ந்த 't' சோதனையாகப் பயன்படுத்தப்படவில்லை என்பது குறிப்பிடத்தக்கதாகும்.

n_1, n_2 ஆகிய அளவுகள் பெரியதாக இருந்தால் 'மேன்-விட்னி *U*' சோதனையானது, ராண்டம் மயமான சோதனைக்கு ஒரு நல்ல தோராயமாக அமைகிறது. மேலும் 'மேன்-விட்னி *U*' சோதனையானது, இந்த 'பிட்மேன்' சோதனையைவிடச் சக்தி வாய்ந்ததாகவும் அமைகிறது என்று கண்டறியப்பட்டுள்ளது: எனவே, '*t*' சோதனைக்கு மாற்றாக, 'மேன்-விட்னி *U*' சோதனையையே பயன்படுத்தலாம்.

பயிற்சி 3

1 'பிஷரி'ன் சரியான நிகழ்தகவு சோதனையைப்பற்றி விரிவாக எடுத்துரைத்து, இச் சோதனைக்கும் கை-வர்க்கச் சோதனைக்குமிடையேயுள்ள வேறுபாட்டைக் கூறுக.

2. குமாஸ்தா தொழில், மற்றத் தொழில்கள் என்று இரு விதமான பணிகளைச் செய்யும் ஆண், பெண் என்ற இருபாலர் களின் அலைவெண்கள் 2 × 2 இணைப்புப் பட்டியலில் கீழ்க்கண்ட வாறு கொடுக்கப்பட்டுள்ளன.

	குமாஸ்தா தொழில்	மற்றத் தொழில்கள்	மொத்தம்
ஆண்கள்	3	9	12
பெண்கள்	3	0	3
மொத்தம்	6	9	15

மேலுள்ள விவரங்களைக் கொண்டு, ஆண்களும் பெண்களும் சமமான விகிதத்தில் தொழில்களை நாடுகின்றனரா என்பதை 'பிஷரி'ன் சரியான நிகழ்தகவுச் சோதனை மூலம் ஆய்க.

3. (அ) 'பிஷரி'ன் சரியான நிகழ்தகவுச் சோதனைக்கு 'டாச்சரி'ன் மாற்று அமைப்பைப் பற்றி விவரித்து எழுதுக.

(ஆ) பின்வரும் 2×2 பட்டியலுக்கு, ஃபிஷரின் சரியான நிகழ்தகவுச் சோதனைக்கான டாச்சரின் மாற்று அமைப்புச் சோதனையைச் செய்து, +, — ஆகிய குறிகளின் விகிதங்களுக்குப் பண்பு செய்யப்பட்ட முதற் குழுவும், இரண்டாம் குழுவும் வேறு பட்டு மிகைத்தன்மை செய்கிறதா என்பதை அறிக.

	குறிகள் −	+	மொத்தம்
குழு 1	3	6	9
குழு 2	5	3	8
மொத்தம்	8	9	17

4. இரு சார்பற்ற கூறுகளுக்கு, சார்பற்ற தன்மையைப் பற்றி அறிய உதவும் கை-வர்க்கச் சோதனையைப்பற்றி விரிவாக எழுதுக. மேலும், இச் சோதனையில் தொடர்ச்சிக்கான 'யேட்ஸ்' திருத்தத்தை எப்பொழுது மேற்கொள்கிறோம் என்பதைக் கூறுக.

5. பச்சை, மஞ்சள், சிவப்பு, வெள்ளை ஆகிய நான்கு வண்ணங்களை விரும்பும் தன்மை பற்றிய ஆய்வு ஒன்றை, 150 நபர்களை ஆண்கள், பெண்கள் என்று கொண்டு நடத்தியபோது, பெற்ற விவரங்கள் கீழ்க்காணும் இணைப்புப் பட்டியலில் குறிக்கப் பட்டுள்ளன.

நபர்கள்	வண்ணங்கள் பச்சை	மஞ்சள்	சிகப்பு	வெள்ளை	மொத்தம்
ஆண்	14	17	21	33	85
பெண்	16	12	15	22	65
மொத்தம்	30	29	36	55	150

வண்ணங்களை விரும்பும் தன்மையில், ஆண்களும் பெண்களும் வேறுபடுகின்றனரா என்பதை 5% மிகைத்தன்மை மட்டத்தில் கை-வர்க்க இருகூறுச் சோதனை மூலம் ஆய்க.

6. கனரக வாகனங்களை ஓட்டும் ஓட்டுநர் போக்குவரத்து நகர்ப்புறப் பேருந்துகளை ஓட்டுவதற்குத் தேர்ந்தெடுக்க, ஆரம்பத் திலும், இறுதியிலும் மேற்கொண்ட முடிவுகள் பின்வருமாறு.

		ஆரம்பச் சோதனை		
		தேறியவர்கள்	தேருதவர்கள்	மொத்தம்
இறுதிச் சோதனை	தேறியவர்கள்	61	35	96
	தேருதவர்கள்	17	7	24
	மொத்தம்	78	42	120

ஓட்டுநர்கள், தேர்ச்சி பெறுவது, அல்லது பெருதது ஆரம்பச் சோதனையையோ அல்லது இறுதிச் சோதனையையோ தழுவி இருக்கிறதா, இல்லையா என்பதை கை-வர்க்க இருகூறுச் சோதனை மூலம் ஆய்க.

7. ஒரு பள்ளியைச் சேர்ந்த 40 மாணவர்களின் கல்வித் திறமை, விளையாட்டுத் திறமை ஆகியவற்றைப் பற்றிய விவரங்கள் கீழ்க்கண்ட 2×2 இணைப்புப் பட்டியலில் கொடுக்கப் பட்டுள்ளன.

		கல்வி		
		திறமையுடையவர்கள்	திறமையற்றவர்கள்	மொத்தம்
விளையாட்டு	திறமையுடையவர்கள்	12	9	21
	திறமையற்றவர்கள்	13	6	19
	மொத்தம்	25	15	40

கல்வித் திறமை, விளையாட்டுத் திறமை ஆகியவை சார் பற்றனவா, இல்லையா என்பதை கை-வர்க்க இருகூறுச் சோதனை மூலம் 5% மிகைத்தன்மை மட்டத்தில் ஆய்க.

8. நுண்ணறிவு, ஆக்கத்திறன், ஆகியவை பற்றி 70 மாணவர்களுக்கு நடத்திய உளவியல் சோதனையில் அவர்கள் மிகுதியாகவும். குறைவாகவும் பெற்ற மதிப்பெண்கள் பின்வரும் (2×2) இணைப்புப்பட்டியலில் குறிக்கப்பட்டுள்ளன.

		நுண்ணறிவு		
		மிகுதியான	குறைவான	மொத்தம்
ஆக்கத்திறன்	மிகுதியான	24	14	38
	குறைவான	12	20	32
	மொத்தம்	36	34	70

மேலுள்ள விவரங்கள் வாயிலாக மிகுதி, குறைவு என்ற இரு மாறிகளும் சார்பற்றனவா என்பதை 5% மிகைத்தன்மை மட்டத்தில் கை-வர்க்க இருகூறுச் சோதனை மூலம் ஆய்க.

9. தாயும், சேய்களும் அவர்களின் மனப்போக்குபற்றிய விவரங்கள் கீழே கொடுக்கப்பட்டுள்ளன.

		தாய்		
		அவசரகுணம்	நிதான குணம்	மொத்தம்
சேய்கள்	அவசர குணம்	55	25	80
	நிதான குணம்	15	75	90
	மொத்தம்	70	100	170

மேலுள்ள விவரங்களைக் கொண்டு, பண்புகளைப் பொறுத்த வரை, தாய், சேய்களின் மனப்போக்கு சார்புடையதா இல்லையா என்பதை கை-வர்க்க இருகூறுச் சோதனை மூலம் ஆய்க.

10. இடைநிலைச் சோதனையின் அமைப்பைப் பற்றி விவரிக்கவும். மேலும், ஃபிஷரின் சரியான நிகழ்தகவுச் சோதனைக்கும் இதற்குமுள்ள முரண்பாடு என்ன?

11. இரு வகைப்பட்ட கரும்பு உற்பத்தியில் முதல் ரகம் 15 எண்ணிக்கையும் இரண்டாம் ரகத்தில் 14 எண்ணிக்கையும் பின்வரும் பலன்களை அளித்தன.

முதல் ரகம்	இரண்டாம் ரகம்
8	6
28	9
10	7
24	18
36	23
15	5
22	25
31	32
11	38
14	26
19	12
21	16
27	34
20	17
13	

முதல் ரகத்திற்கான இடைநிலைக்கும், இரண்டாம் ரகத்திற்கான இடைநிலைக்கும் இடையே வேறுபாடு இருக்கிறதா, இல்லையா என்பதை இடைநிலைச் சோதனை மூலம் ஆய்க.

12. 12 ஆண்களும், 20 பெண்களும் ஒரு வினாத்தாளுக்கு விடையளிக்கையில் கீழ்க்கண்ட மதிப்பெண்களைப் பெற்றனர். விடையளித்து மதிப்பெண்களைப் பெறுவதில் ஆண்களும், பெண்களும் வேறுபடுகின்றனரா என்பதை இடைநிலைச் சோதனையைப் பயன்படுத்தி ஆய்க.

ஆண்களின் மதிப்பெண்கள்	பெண்களின் மதிப்பெண்கள்
30	43
21	43
39	35
47	37
46	40
47	34
48	31
49	43
51	43
60	45
38	22
37	23
	25
	36
	37
	42
	54
	56
	48

13. ஆராய்ச்சித் துறையில் ஈடுபட்டுள்ள மாணவர்களில் 5 மாணவர்கள் தாய்நாட்டைச் சேர்ந்தவர்களாகவும், 4 மாணவர்கள் வெளிநாட்டைச் சேர்ந்தவர்களாகவும் இருக்கின்றனர். இம் மாணவர்களின் ஆளுமைத்திறன் (personality ability) பற்றிய மதிப்புகள் கீழுள்ள அட்டவணையில் கொடுக்கப்பட்டுள்ளன.

X: தாய்நாட்டு மாணவர்கள்	*Y*: வெளிநாட்டு மாணவர்கள்
59	98
62	79
64	45
45	58
53	

வெளிநாட்டு மாணவர்கள், தாய்நாட்டு மாணவர்களைவிட ஆளுமைத்திறனை அதிகமாகப் பெற்றுள்ளனரா என்பதை மேன்-விட்னி *U* சோதனையைப் பயன்படுத்தி ஆய்க.

14. ஒரு வகுப்பறையிலுள்ள மாணவர்கள், ராண்டம் முறையில் இரு பிரிவுகளாகப் பிரிக்கப்பட்டனர். ஒரு பிரிவிலுள்ள மாணவர்களுக்கு ஆசிரியரின் சொற்படி சீரான முறையில் படிக்கக் கற்றுக்கொடுக்கப்பட்டது. மற்றொரு பிரிவிலுள்ள மாணவர்களுக்கு அவர்களாகவே சுயேச்சையாக ஓர் ஆசிரியரின் மேற்பார்வையில் படிக்கக் கற்றுக்கொண்டனர். ஆண்டு இறுதியில் படிக்கும் திறனுக்கான தேர்வு நடத்தப்பட்டதில் இவ்விரு பிரிவு மாணவர்கள் பெற்ற மதிப்பெண்கள் பின்வருமாறு:

முதற்பிரிவு	இரண்டாம் பிரிவு
35	36
48	42
63	55
56	37
72	46
46	49
54	52
58	44
65	62
67	39
66	41
47	47
38	59
	57
	45
	31
	44
	34
	40
	51
	42

மேலுள்ள விவரங்களுக்கு மேன்-விட்னி *U* சோதனையைப் பயன்படுத்தி இவ்விரு சார்பற்ற பிரிவுகளின் சராசரிகளுக்குள் வேறுபாடு உள்ளதா என்பதை ஆய்க.

15. கோல்மோகோராவ்-ஸ்மிர்னுவின் இருகூறுச் சோதனையில் பயன்படுத்தப்படும் எடுகோள்களைக் குறிப்பிட்டு இச்சோதனையைப் பற்றி விளக்குக.

16. A, B என்ற இரு வெவ்வேறு பல்கலைக் கழகங்களில் இருந்து, முறையே $X_1, X_2, \ldots X_{10}; Y_1, Y_2 \ldots Y_{17}$ என்ற மாணவர்கள் ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டனர். கல்வித் திறனை அறிய இவ்விரு குழுக்களுக்கும் பொதுவான தேர்வு நடத்தப்பட்டதில், அம் மாணவர்கள் பெற்ற மதிப்பெண்கள் பின்வருமாறு:

A பல்கலைக்கழகம்	B பல்கலைக்கழகம்
7·5	5·4
8·6	5·7
8·7	5·9
9·3	6·6
9·4	6·8
9·8	8·2
10·1	9·1
10·5	9·9
11·5	10·8
12·6	11·3
	11·5
	11·6
	11·8
	12·1
	12·4
	12·7
	13·1

இரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட இவ்விரு கூறுகளுக்கிடையே, கல்வித் திறனில் வேறுபாடு உள்ளதா என்பதை, கோல்மோகோராவ்-ஸ்மிர்னுவின் சோதனையைப் பயன் படுத்தி ஆய்க.

17. க்ராமர்–வான் மைஸஸின் இருகூறுச் சோதனையைப் பற்றி விவரித்து, இதற்கும், கோல்மோகோராவ்-ஸ்மிர்னுவின் இரு கூறுச் சோதனைக்கும் இடையேயுள்ள தொடர்பை விவரிக்கவும்.

18. (16)-லுள்ள விவரங்களுக்கு க்ராமர்-வான் மைஸஸ் இருகூறுச் சோதனையைச் செய்க.

19. பேச்சுத் திறனை அறியும்பொருட்டு ஒரு நகரத்திலிருந்து 6 மாணவர்களையும், மற்றொரு நகரத்திலிருந்து 9 மாணவர்

களையும் தேர்ந்தெடுத்து, பேச்சுத் திறனுக்கான மதிப்பெண்கள் கொடுக்கப்பட்டுக் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப் பட்டது :

நகரம் 1	நகரம் 2
65	56
49	58
32	35
46	39
84	57
48	63
	72
	81
	92

பேச்சுத் திறனில் இரு நகர்ப்புறத்து மாணவர்களிடையே ஏதாவது வேறுபாடு உள்ளதா என்பதை க்ராமர்-வான் மைஸஸின் சோதனையைப் பயன்படுத்தி ஆய்க.

20. (அ) வால்டு-வுல்ஃபாவிட்ஸின் ஓட்டச் சோதனையைப் பற்றி விவரிக்கவும்.

(ஆ) 5 வயதுடைய 10 சிறுவர்களும், சிறுமிகளும் சேர்ந்து விளையாடும் போக்கில் ஆக்கிரமிப்புத் தன்மை இருபாலாரிடத்தும் ஒரே மாதிரியாக உள்ளதா, இல்லையா என்பதை அறிய அவர் களின் ஆக்கிரமிப்பு மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டன. இவ் விவரங்களைக் கொண்டு வால்டு-வுல் ஃபாவிட்ஸ் சோதனையைப் பயன்படுத்திச் சிறுவர், சிறுமியர் களிடையே ஆக்கிரமிப்புத் தன்மை ஒரே மாதிரியாக உள்ளதா இல்லையா என்பதை ஆய்க:

சிறுவர்கள்	சிறுமியர்கள்
66	78
59	49
84	32
92	68
120	25
75	15
125	24
59	26
148	37
112	29

21. சாதாரண நிலையில் பயிற்சி பெற்ற 22 குதிரைகள் ஓடிய தூரமும் (கிலோமீட்டரில்), கடல் மணலில் ஓடிப் பயிற்சி பெற்ற 9 குதிரைகள் ஓடிய தூரமும் (கிலோமீட்டரில்) பின்வரும் அட்டவணையில் குறிக்கப்பட்டுள்ளன:

சாதாரண நிலையில் பயிற்சி பெற்ற குதிரைகள்	கடல் மணலில் ஓடிப் பயிற்சி பெற்ற குதிரைகள்
22	28
18	31
24	29
25	27
26	28
21	34
27	35
17	21
28	20
26	
30	
19	
23	
33	
30	
20	
17	
16	
23	
29	
18	
16	

மேலுள்ள விவரங்களைக் கொண்டு சாதாரண நிலையில் பயிற்சி பெற்ற குதிரைகள் ஓடும் தூரத்திற்கும், கடல் மணலில் ஓடிப் பயிற்சி பெற்ற குதிரைகள் ஓடும் தூரத்திற்கும் இடையே வேறுபாடு உள்ளதா என்பதை வால்டு-வுல்ஃபாவிட்ஸ் சோதனையைப் பயன்படுத்தி ஆய்க.

22. மோஸஸின் சோதனையை எம்மாதிரியான விவரங்களுக்குப் பயன்படுத்துகிறோம் என்பதைக் கூறி அச்சோதனையின் அமைப்பை விவரிக்கவும்.

23. வேடிக்கைக் (circus) காட்சிகளில் வரும் மெய்சிலிர்க்கும் நிகழ்ச்சிகளைப் பார்க்கும் இரு குழுக்களின் ஆளுமைச் சோதனை மதிப்புகளை ஒப்பிட்டுப் பார்த்ததில் *X* என்ற குழுவில் உள்ள வர்கள் பயமடையாத நிலையிலும், *Y* என்ற குழுவில் உள்ளவர்கள் பயமடைந்த நிலையிலும் இருப்பதாக அறியப்பட்டது. பிறகு *X* குழுவிலுள்ள 11 நபர்களுக்கும், *Y* குழுவிலுள்ள 11 நபர்களுக்கும் மெய்சிலிர்க்கும் வேடிக்கைக் காட்சிகள் காட்டப்பட்டு அவர்களின் ஆளுமை மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப் பட்டன:

X	Y
14	18
18	25
9	5
12	16
12	0
4	22
10	15
10	9
8	8
5	8
15	14

X குழுவிலுள்ள பயமடையாதவர்களின் ஆளுமைத் திற னுக்கும், *Y* குழுவிலுள்ள பயமடைபவர்களின் ஆளுமைத் திறனுக்கும் வேறுபாடு உள்ளதா இல்லையா என்பதை 'மோஸஸ்' சோதனையைப் பயன்படுத்தி ஆய்க.

24. இரு சார்பற்ற கூறுகளுக்கான, ராண்டம் மயமான சோதனையைப் பற்றி விவரித்து எழுதவும். மேலும் இச் சோதனைக்கும், பிட்மேன் சோதனைக்கும் இடையேயுள்ள தொடர்பு என்ன?

4. k தொடர்புடைய கூறுகளுக்கான சோதனைகள்

(Tests for *k* Related Samples)

இதற்கு முந்தின அத்தியாயங்களில், ஒருகூறுச் சோதனைகளை யும், சார்புடைய அல்லது சார்பற்ற இருகூறுகளுக்கான புள்ளியியல் சோதனைகளையும் கண்டோம். இவ்வத்தியாயத்தில், மூன்றும் அதற்கும் மேற்பட்ட கூறுகளுக்கான புள்ளியியல் சோதனைகளைப் பற்றிக் காண்போம்.

இருகூறுச் சோதனையையே, 4 கூறுகளைக் கொண்ட விவரங் களுக்குப் பயன்படுத்த விரும்பினால், $\begin{pmatrix} 4 \\ 2 \end{pmatrix} = 6$ என அமை வதால் நம்முடைய கண்டறிந்த கூறுகளில் ஒன்று இந்த 6 புறக் கோடியாக அமைந்த சாத்தியமான வெளிப்பாடுகளிலுள்ளதால் நம்முடைய கண்டறிந்தவைகளைக் கொண்டு H_0-ஐ $\mu_X \neq \mu_Y$ என்ற மாற்று எடுகோளுக்குச் சாதகமாக மறுத்திருக்க முடியும். H_0-ன் கீழ் இருமுனைச் சோதனைக்கு அத்துமீறிய சாத்தியமான வெளிப்பாடு களில் ஒன்று நம்முடைய கண்டறிந்த மதிப்பு இருப்பதற்கான சரியான நிகழ்தகவு $p = \frac{6}{126} = 0{\cdot}048$ ஆகும்.

குறிப்பு : n_1, n_2 ஆகிய அளவைகள் மிகவும் பெரியதாக இருந்தால் ராண்டம் மயமான சோதனையைப் பயன்படுத்துவது மிகவும் கடினமாகும். ஆனால் n_1, n_2 ஆகிய அளவுகள் பெரிய தாகவும் அக் கூறுகளின் இணைந்த தொடரின் தட்டை அளவு (Kurtosis) சிறியதாகவும், n_1, n_2 ஆகியவைகளின் விகிதம் (Ratio) $\frac{1}{5}$ லிருந்து 5 வரை இருந்தால் $\begin{pmatrix} n_1 + n_2 \\ n_2 \end{pmatrix}$ சாத்தியமான வெளிப்பாடுகளுள்ள ராண்டம் மயமான பரவலானது தோராய மாக t பரவலானது அமைகிறது என்று பிட்மேன் என்பவர் மாற்றுச் சோதனையைக் கண்டுள்ளார்.

மேலுள்ள நிபந்தனைகளுக்கு இணங்கினால்,

$$t = \frac{\dfrac{\Sigma X}{n_X} - \dfrac{\Sigma Y}{n_Y}}{\sqrt{\dfrac{\Sigma\left(Y - \dfrac{\Sigma Y}{n_2}\right)^2 + \Sigma\left(X - \dfrac{\Sigma X}{n_1}\right)^2}{n_X + n_Y - 2}\left(\dfrac{1}{n_X} + \dfrac{1}{n_Y}\right)}}$$

என்று $n_X - n_Y - 2$ வரையற்ற பாகையைப் பெற்ற தோராய *t* பரவலாக அமைகிறது. இந்த *t* சோதனைக்கான இரு கூறுகளை அமைத்து ஒவ்வொரு ஜோடிக்கும் சோதனை செய்ய வேண்டியுள்ளது. இம்மாதிரியான முறை கடினமானது என்பது மட்டுமன்றி தவறான முடிவுகளையும் அளிக்கும். $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதி $k = 4$ கூறுகளுக்குள் வேற்றுமை இல்லை என்ற சூனிய எடுகோளுக்கு 6 இருகூறுச் சோதனைகளைச் செய்வதால் H_0-ஐ ஏற்றுக்கொள்ளவோ மறுக்கவோ செய்யும் தீர்மானத்திற்கு ஒரு வாய்ப்புக் கொடுப்பதற்குப் பதிலாக 6 வாய்ப்புகளை அளிக்கிறோம்.

இதனால், $\alpha = 0{\cdot}05$ மிகைத்தன்மை மட்டத்தில் தவறுதலாக H_0-ஐ மறுக்கும் நிலையை ஏற்படுத்துகிறோம். அதாவது பிழை I வகையை உண்டாக்குகிறோம். எனவே, இம்மாதிரியான தவறுதல்களைத் தவிர்த்துச் சோதனையை மேற்கொள்வதற்கு *k* கூறுகளுக்கான தனிச் சோதனைகள் உள்ளன. அவை

(i) கோச்ரான் *Q* சோதனை

(ii) ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வுச் சோதனைகள் ஆகும்.

'கோச்ரான் *Q*' சோதனை (Cochran *Q* Test)

இரு சார்புடைய கூறுகளுக்குப் பயன்படுத்தப்படும் மெக்னமார் சோதனையின் நீட்டிப்பே கோச்ரான் *Q* சோதனையாகும். இச் சோதனை மூன்று அல்லது அதற்கு மேற்பட்ட ஜோடியான கூறுகளுக்குள் வேற்றுமை இருக்கிறதா இல்லையா என்பதை அறிய உதவுகிறது. கொடுக்கப்பட்ட *k* ஜோடிப் பொருத்தமான கூறுகளை *N* நிரைகளையும், *k* நிரல்களையும் பெற்ற விவரமாகக் கொண்டு ஒவ்வொரு நிரலிலுமுள்ள பிரதிச்செயலின் (response) விகிதம் அல்லது அலைவெண் ஒரே மாதிரியாக இருக்கும் என்ற சூனிய எடுகோளைச் சோதனை செய்கிறோம். சூனிய

எடுகோள் உண்மையானல், அதாவது ஒவ்வொரு நிலையிலும் 'வெற்றி' நிகழ்தற்கான நிகழ்தலில் வித்தியாசமில்லை என்றிருந்து நிரைகளின் எண்ணிக்கை மிகச் சிறியதாக இல்லாமல் இருந்தால்,

$$Q = \left\{ \frac{k(k-1) \sum_{j=1}^{k} (G_j - \bar{G})^2}{k \sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2} \right\}$$

ஆனது $(k-1)$ வரையற்ற பாகையைப் பெற்ற தோராயமான கைவர்க்கப் பரவலாக அமைந்துள்ளது என்று கோச்ரான் நிரூபித்துள்ளார்.

இதில் $G_j = j^{th}$ நிரலிலுள்ள வெற்றிகளின் மொத்த எண்ணிக்கையாகும்.

$\bar{G} = G_j$-ன் சராசரியாகும்.

$L_i = i$ ஆம் நிரையிலுள்ள வெற்றிகளின் மொத்த எண்ணிக்கையாகும்.

மேலுள்ள சூத்திரத்தை எளிய முறையில்,

$$Q = \left\{ \frac{(k-1)\left[k \sum_{j=1}^{k} G_j^2 - \left(\sum_{j=1}^{k} G_j \right)^2 \right]}{k \sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2} \right\}$$

என்று எழுதலாம். இது $(k-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. Q-ன் கண்டறிந்த மதிப்பானது, $(k-1)$ வரையற்ற பாகைக்கும் α என்ற மிகைத் தன்மை மட்டத்திற்கும் ஏற்ற அட்டவணை C-யிலுள்ள மதிப்புகளுக்குச் சமமாக அல்லது அதிகமாக இருந்தால் H_0-ஐ மறுக்கிறோம்.

மாதிரி : 18 குடும்பப் பெண்மணிகள் மூன்று பேட்டி காண்பவர்களின் வினவுக்கு அளித்த 'ஆம்—இல்லை' விடைகள்

'ஆம்' என்றால் 1 என்றும், 'இல்லை' என்றால் 0 என்றும் குறிப்பிட்டுப் பின்வருமாறு அட்டவணைப்படுத்தப்பட்டன:

குடும்பப் பெண்மணிகள்	முதல் பேட்டி காண்பவரின் கேள்விக்கான பதில்	இரண்டாம் பேட்டி காண்பவரின் கேள்விக்கான பதில்	மூன்றாம் பேட்டி காண்பவரின் கேள்விக்கான பதில்
1	1	1	0
2	1	1	0
3	1	1	1
4	1	1	0
5	1	1	0
6	1	1	0
7	1	1	1
8	1	1	1
9	0	0	0
10	1	0	0
11	0	1	0
12	1	1	0
13	1	1	0
14	1	0	0
15	0	0	0
16	0	1	0
17	1	1	0
18	0	0	0

மூன்று பேட்டி காண்பவர்களுக்கு அளிக்கப்படும் 'ஆம்' என்ற பதிலுக்கான நிகழ்தகவு சமமானதே என்ற சூனிய எடுகோளையும், மூன்று பேட்டி காண்பவர்களுக்கு அளிக்கப்படும் 'ஆம்' என்ற பதிலுக்கான நிகழ்தகவுகள் வேறுபட்டுள்ளன என்ற மாற்று

எடுகோளையும் கொண்டு கோச்ரான் Q சோதனை செய்வோம். $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம்.

குடும்பப் பெண்கள்	முதல் பேட்டி	இரண்டாம் பேட்டி	மூன்றும் பேட்டி	L_i	L_i^2
1	1	1	0	2	4
2	1	1	0	2	4
3	1	1	1	3	9
4	1	1	0	2	4
5	1	1	0	2	4
6	1	1	0	2	4
7	1	1	1	3	9
8	1	1	1	3	9
9	0	0	0	0	0
10	1	0	0	1	1
11	0	1	0	1	1
12	1	1	0	2	4
13	1	1	0	2	4
14	1	0	0	1	1
15	0	0	0	0	0
16	0	1	0	1	1
17	1	1	0	2	4
18	0	0	0	0	0
	$G_1 = 13$	$G_2 = 13$	$G_3 = 3$	$\sum_{i=1}^{18} L_i = 29$	$\sum_{i=1}^{18} L_i^2 = 63$

$G_1 = 13, \quad G_2 = 13, \quad G^3 = 3$

$$\sum_{i=1}^{18} L_i = 29 \quad \text{and} \quad \sum_{i=1}^{18} L_i^2 = 63$$

$$\sum_{j=1}^{3} G_j = 13 + 13 + 3 = 29$$ ஆகும்.

எனவே,

$$Q = \frac{(k-1)\left[k\sum_{j=1}^{k} G_j^2 - \left(\sum_{j=1}^{k} G_j\right)^2\right]}{k\sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2}$$

$$= \frac{(3-1)[3\{13^2+13^2+3^2\}-(29)^2]}{3(29)-63}$$

$$= \frac{2[3\{347\}-841]}{87-63}$$

$$= \frac{2[200]}{24} = \frac{400}{24} = 16{\cdot}7$$

அட்டவணை C-லிருந்து H_0-ன்கீழ் $k-1 = 3-1 = 2$ வரையற்ற பாகைக்கு $\alpha = 0{\cdot}01$ மிகைத்தன்மை மட்டத்திற்கான மதிப்பு $9{\cdot}21$ ஆகும். கண்டறிந்த Q-ன் மதிப்பான $16{\cdot}7$ ஆனது $9{\cdot}21$-ஐவிட அதிகமாக உள்ளதால் H_0-ஐ $\alpha = 0{\cdot}01$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனை (Friedman Two-way Analysis of Variance Rank Test)

k கூறுகளானவை ஒரே முழுமைத் தொகுதியிலிருந்து வரப் பெற்றவை என்ற சூனிய எடுகோளைச் சோதனை செய்வதற்கு ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனை பயன் படுகிறது. k கூறுகளும் ஜோடி செய்யப்படுவதால் ஒவ்வொரு கூறிலும் உள்ள அளவுகள் சமமாகவே இருக்கும். N கூறுகள் k நிலைகளில் வேறுபடுகின்றனவா என்பதை அறியலாம். உதாரண மாக, 3 மாணவர்களுக்கு 4 விதமான கற்பிக்கும் முறைகள் கையாளப்பட்டன எனில் கற்பிக்கும் முறைகளில் வேறுபாடு உள்ளதா என்பதை இச் சோதனையின் மூலம் அறியலாம். இங்கு $N = 3$, $k = 4$ ஆகும்.

ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனையைச் செய்ய N நிரைகளும், k நிரல்களும் கொண்ட இருவழி அட்டவணையாக அமைந்த விவரங்கள் தேவைப்படுகின்றன. இங்கு நிரைகள் பல்வேறு கூறுகளையும், நிரல்கள் பல்வேறு நிலைகளையும் குறிக்கின்றன.

இச் சோதனை ஒரு தரச் சோதனையாகும். ஒவ்வொரு நிரையிலுள்ள மதிப்புக்களும் தரப்படுத்தப்படுகின்றன. தரங்களைக் கொண்ட ஒவ்வொரு நிரலும் ஒரே முழுமைத் தொகுதியில் இருந்து வரப்பெற்றவையா என்பதை இச் சோதனை ஆய்வு செய்கிறது. உதாரணமாக 3 கூறுகள், 4 நிலைகளில் கண்ட மதிப்புகளைக் கீழ்க்கண்ட இருவழி அட்டவணையில் அமைப்போம்.

	நிலைகள்			
	I	II	III	IV
கூறு I	11	6	3	9
கூறு II	8	7	4	10
கூறு III	11	3	4	8

இங்கு $k=4$, $N = 3$ ஆகும்.

இப்பொழுது ஒவ்வொரு நிரையின் மதிப்புகளையும் தரப்படுத்திக் கீழ்க்கண்டவாறு அமைக்கிறோம்.

தரப்படுத்தப்பட்ட இருவழி அட்டவணை

	நிலைகள்			
	I	II	III	IV
கூறு I	4	2	1	3
கூறு II	3	2	1	4
கூறு III	4	1	2	3
R_j	11	5	4	10

H_0 உண்மை எனில் ஒவ்வொரு நிரலிலுமுள்ள தரங்களின் பரவல் தற்செயலாக ஏற்பட்டது என்று கொள்கிறோம். எனவே 1, 2, 3, 4 என்ற தரங்கள் எல்லா நிரல்களிலும் ஓரளவு சமமான அலைவெண்ணோடு நிகழும் என்று எதிர்பார்க்கிறோம். இச் சோதனையைச் செய்ய χ_r^2 என்ற புள்ளியியல் அளவையின் மதிப்பைக் காணவேண்டும். நிரைகளின் அல்லது நிரல்களின் எண்ணிக்கை மிகச் சிறியதாக இல்லாமலிருந்தால் χ_r^2 ஆனது $(k-1)$ வரையற்ற பாகையைப் பெற்ற தோராயமான கை-வர்க்கப் பரவலாக அமைகிறது.

$$\chi_r^2 = \left\{\frac{12}{N\,k\,(k+1)} \sum_{j=1}^{k} (R_j)^2\right\} - 3\,N\,(k+1)$$

இங்கு k = நிரல்களின் எண்ணிக்கை ஆகும்.

N = நிரைகளின் எண்ணிக்கை ஆகும்.

R_j = jஆம் நிரலிலுள்ள தரங்களின் கூட்டல் ஆகும்.

χ_r^2-ன் கண்டறிந்த மதிப்பானது α என்ற மிகைத்தன்மை மட்டத்திற்கு $(k-1)$ வரையற்ற பாகைக்கு ஏற்ற அட்டவணை C-யிலுள்ள மதிப்பிற்குச் சமமாகவோ குறைவாகவோ இருந்தால் வெவ்வேறு நிரலின் தரங்களின் கூட்டல்களுக்குள் மிகைத்தன்மையான வேறுபாடு உள்ளது என்று உய்த்துணருகிறோம். அதாவது, H_0-ஐ மேலுள்ள உதாரணத்திற்கு $N = 3$, $k = 4$, $R_1 = 11$, $R_2 = 5$, $R_3 = 4$, $R_4 = 10$ ஆகும். எனவே,

$$\begin{aligned}
\chi_r^2 &= \frac{12}{(3)(4)(5)}\left\{(11)^2 + (5)^2 + (4)^2 + (10)^2\right\} - 3\,(3)\,(5)\\
&= \frac{12}{60}\left\{(11)^2 + (5)^2 + (4)^2 + (10)^2\right\} - 3\,(3)\,(5)\\
&= \frac{1}{5}\left\{121 + 25 + 16 + 100\right\} - 3\,(3)\,(5)\\
&= \frac{1}{5}\left\{121 + 25 + 16 + 100\right\} - 45\\
&= \frac{1}{5}\left\{262\right\} - 45\\
&= 52{\cdot}4 - 45\\
&= 7{\cdot}4 \text{ ஆகும்.}
\end{aligned}$$

அட்டவணை C-யிலிருந்து $k-1 = 4-1 = 3$ வரையற்ற பாகைக்கு $0 \cdot 01$ மிகைத்தன்மை மட்டத்திற்கான மதிப்பு $11 \cdot 34$ ஆகும். கண்டறிந்த χ_r^2-ன் மதிப்பு $7 \cdot 4$ ஆகும். இது அட்டவணையின் மதிப்பான $11 \cdot 34$-ஐவிடக் குறைவாக உள்ளதால் H_0-ஐ $\alpha = 0 \cdot 01$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். அதாவது 4 கூறுகளும் ஒரே முழுமைத் தொகுதியிலிருந்து வரப் பெற்றவை என்ற சூனிய எடுகோளை $\alpha = 0 \cdot 01$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். மேலும், தெளிவாக R_j-களுக்குள் வேற்றுமையுள்ளதா என்று அறிகிறோம்.

மாதிரி : மூன்றுவிதமான நிலைகளில் ஒரு குழுவிலுள்ள 12 நபர்கள் செய்த பிரதிச் செயல்களின் (responses) மதிப்புகள் பின்வருமாறு:

குழு	நிலைகள்		
	1	2	3
1	6	10	8
2	7	9	5
3	4	7	6
4	3	5	4
5	2	2	4
6	3	4	6
7	7	6	5
8	9	8	7
9	5	3	2
10	4	6	7
11			
12			

மூன்று நிலைகளுக்குள் வேற்றுமையில்லை என்ற சூனிய எடு கோளையும் வேற்றுமையுள்ளது என்ற மாற்று எடுகோளையும் கொண்டு ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனை செய்வோம். முதலில் ஒவ்வொரு குழுவிற்கும் தரங்களை அளித்துப் பின்வரும் அட்டவணை அமைக்கிறோம்:

தரப்படுத்தப்பட்ட இருவழி அட்டவணை

குழு	நிலைகள்		
	1	2	3
1	1	3	2
2	2	3	1
3	1	3	2
4	1	3	2
5	1·5	1·5	3
6	1	2	3
7	3	2	1
8	3	2	1
9	3	2	1
10	1	2	3
R_j	17·5	23·5	19

பின்பு $$\chi_r^2 = \left\{\frac{12}{Nk(k+1)}\sum_{j=1}^{k}(R_j)^2\right\} - 3N(k+1)$$

$$= \left\{\frac{12}{(10)(3)(3+1)}\left[(17\cdot5)^2+(23\cdot5)^2+(19)^2\right] - 3(10)(3+1)\right\}$$

$$= \left\{\frac{12}{120}(306\cdot25+552\cdot25+361) - 120\right\}$$

$$= \left\{\frac{1}{10}(1219\cdot5) - 120\right\}$$

$$= 121\cdot95 - 120$$

$$= 1\cdot95$$

அட்டவணை C-யிலிருந்து $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்திற்கு $k-1 = 3-1 = 2$ வரையற்ற பாகைக்கான χ_r^2-ன் மதிப்பு 5·99 ஆகும். கண்டறிந்த χ_r^2-ன் மதிப்பான 1·95 ஆனது அட்டவணையின் மதிப்பைவிடக் குறைவாக இருப்பதால் H_0-ஐ $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்தில் ஏற்றுக்கொள்கிறோம்.

பயிற்சி 4

1. k தொடர்புடைய கூறுகளுக்கான கோச்ரான் Q சோதனையைப் பற்றி விவரிக்கவும்.

2. ஓர் உளவியல் சோதனையில் 19 மாணவர்கள் X, Y, Z என்ற மூன்று வெவ்வேறு உளவியல் நிபுணர்களின் கேள்விகளுக்கு அளித்த 'ஆம்–இல்லை' விடைகளின் விவரங்கள் பின்வருமாறு:

மாணவர்கள்	உளவியல் நிபுணர்கள்		
	X	Y	Z
a	1	0	1
b	0	1	1
c	1	1	0
d	0	1	1
e	1	1	0
f	0	0	1
g	1	0	1
h	1	1	0
i	0	0	1
j	1	1	0
k	0	1	1
l	0	0	0
m	1	0	1
n	1	1	0
o	1	0	0
p	1	1	1
q	1	0	0
r	1	1	0
s	1	0	0

மூன்று உளவியல் நிபுணர்களுக்கு அளிக்கப்படும் 'ஆம்-இல்லை' என்ற பதில்களுக்கான நிகழ்தகவில் வேறுபாடு உள்ளதா இல்லையா என்பதை 'கோச்ரான் Q' சோதனை மூலம் ஆய்க.

3. ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனையில் பயன்படுத்தப்படும் எடுகோள்களைக் கூறி அச் சோதனையைப்பற்றி விவரிக்கவும்.

4. ஒரு குழுவிலுள்ள 13 நபர்கள் மூன்று வெவ்வேறு நிலைகளில் செய்த பிரதிச் செயல்களின் மதிப்புகள் பின்வருமாறு :

குழு	நிலைகள் (1)	(2)	(3)
1	5	4	2
2	6	3	4
3	4	2	3
4	3	6	5
5	3	7	7
6	2	5	8
7	7	4	9
8	6	8	6
9	8	7	3
10	9	6	2
11	6	4	2
12	7	3	5
13	5	2	4

மேலுள்ள விவரங்களைப் பயன்படுத்தி மூன்று நிலைகளுக்குள் வேறுபாடு உள்ளதா இல்லையா என்பதை ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனையைப் பயன்படுத்தி ஆய்க.

5. ஒரு கிரிக்கெட் ஆட்டக் குழுவிலுள்ள 11 விளையாட்டு வீரர்கள், ஓவல், அடிலெய்டு, சேப்பாக், ட்ரினிடாட் ஆகிய விளையாட்டு மைதானங்களில் முதல் இன்னிங்ஸில் பெற்ற ரன்கள் பின்வருமாறு :

விளையாட்டு வீரர்கள்	விளையாட்டு மைதானங்கள்: ஓவல்	அடியெய்டு	சேப்பாக்	ட்ரினிடாட்
1	42	15	82	45
2	13	63	105	56
3	12	0	68	75
4	11	1	15	92
5	56	2	1	2
6	75	5	0	0
7	82	11	59	0
8	0	13	173	1
9	12	51	63	2
10	2	102	12	5
11	10	68	11	13

மேலுள்ள விவரங்களிலிருந்து 4 விளையாட்டு மைதானங்களில் விளையாடி பெற்ற ரன்களுக்குள் வேறுபாடு உள்ளதா இல்லையா என்பதை ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனை மூலம் ஆய்க.

5. k சார்பற்ற கூறுகளுக்கான சோதனைகள்

(Tests for k Independent Samples)

பல சார்பற்ற கூறுகள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றனவா இல்லையா என்பதை அறியச் சில சோதனைகள் செய்வதன் மூலம் அறியலாம். இவ்வத்தியாயத்தில் மூன்றும் அதற்கு மேற்பட்ட சார்பற்ற குழுக்கள் அல்லது கூறுகளுக் கிடையே வேறுபாடுகள் உள்ளனவா இல்லையா என்பதை அறியத் தேவைப்படும் சோதனைகளைப் பற்றிக் காண்போம். சாதாரண மாகச் சுட்டுறுப்பைச் சார்ந்த புள்ளியியல் முறைகளில் பல சார்பற்ற கூறுகள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப் பெற்றனவா இல்லையா என்பதை அறிய ஒருவழி பரவற்படி ஆய்வில் (one-way analysis of variance) F சோதனை மூலம் அறிகிறோம். இந்த F சோதனையை மேற்கொள்வதற்கு முன்பு கண்டறிந் தவை சார்பற்ற முறையில் இயல்நிலைப் பரவலைப் பெற்ற முழுமைத் தொகுதிகளிலிருந்து எடுக்கப்பட்டவை என்ற தற்கோளைக் கொள்கிறோம். ஆனால், இம்மாதிரியான தற் கோள்கள் சில விவரங்களுக்கு ஏற்புடையதாக இல்லாமலோ அல்லது உண்மையற்றதாகவோ இருந்தால் சுட்டுறுப்பைச் சாராத k சார்பற்ற கூறுகளுக்கான புள்ளியியல் சோதனைகளை மேற்கொள்ளவேண்டியநிலை ஏற்படுகிறது. மேலும் கொடுக்கப் படும் சில விவரங்கள் பகுக்கப்பட்ட வகையைச் (classificatory) சேர்ந்ததாகவோ அல்லது கண்டறிந்தவையானவை தரங்களைப் (ranks) பெற்றதாகவோ இருந்தால் சுட்டுறுப்பைச் சாராத சோதனைகளைத் தவிர வேறு எந்தச் சோதனைகளையும் மேற்கொள்ள முடியாது. இவ்வத்தியாயத்தில் k சார்பற்ற கூறுகளுக்கான சில சுட்டுறுப்பைச் சாராத சோதனைகளைப் பற்றிக் காண்போம்.

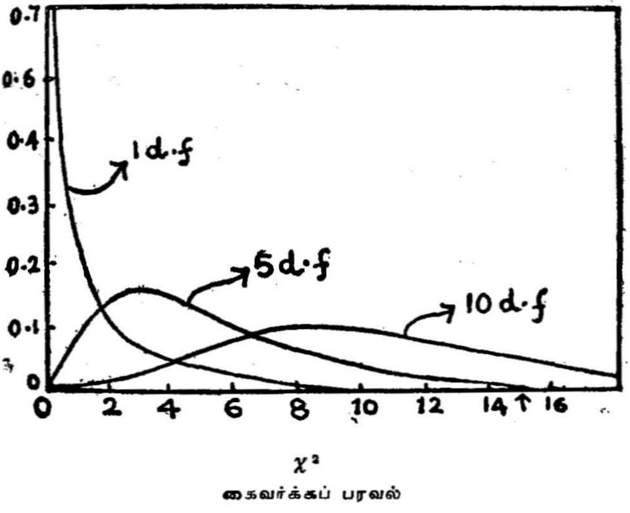
k சார்பற்ற கூறுகளுக்கு கைவர்க்கச் சோதனை (Chi-square Test for k Independent Samples)

கொடுக்கப்பட்டுள்ள விவரங்கள் பகுக்கப்பட்ட வகையைச் சேர்ந்ததாகவோ அல்லது கண்டறிந்தவை தரப்படுத்தப்

பட்டதாக இருக்குமானால் *k* சார்பற்ற கூறுகளுக்கு கைவர்க்கச் சோதனையை அக் கூறுகளுக்குள் வேறுபாடு உள்ளதா இல்லையா என்பதை அறியச் சோதனை செய்கிறோம். இச் சோதனையைப் பயன்படுத்தக் கொடுக்கப்பட்ட விவரத்திலுள்ள அலைவெண்களை $(k \times r)$ அட்டவணையாக அமைக்கிறோம். *k* கூறுகளின் அலைவெண்கள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளைக் கொள்கிறோம். இந்த எடுகோளைச் சோதனை செய்யக் கீழ்க்காணும் சூத்திரத்தைப் பயன்படுத்துகிறோம்.

$$\text{அதாவது } \chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

இதில் O_{ij} என்பது iஆம் நிரையிலும், jஆம் நிரலிலுமுள்ள கண்டறிந்த அலைவெண்ணாகும் (observed frequencies). E_{ij} என்பது H_0-ன் கீழ் எதிர்பார்க்கப்படும் iஆம் நிரையிலும் jஆம் நிரலிலுமுள்ள எதிர்பார்க்கும் அலைவெண்ணாகும். H_0-ன் கீழ் மேலுள்ள சூத்திரத்திலிருந்து காணப்படும் χ^2-ன் மாதிரிப் பரவலானது $(k-1)$ $(r-1)$ வரையற்ற பாகையைப் (degrees of freedom) பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. இங்கு



கைவர்க்கப் பரவல்

k என்பது நிரல்களின் (columns) எண்ணிக்கையையும் *r* என்பது நிரைகளின் (rows) எண்ணிக்கையையும் குறிக்கிறது. கொடுக்கப்

பட்ட விவரத்திற்கு மேலுள்ள χ^2 சூத்திரத்தைப் பயன்படுத்திக் கண்ட மதிப்பானது $(k-1)$ $(r-1)$ வரையற்ற பாகைக்கும், α என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற அட்டவணை C-யிலுள்ள மதிப்புக்குச் சமமாக அல்லது அதிகமாக இருந்தால் α என்ற மிகைத்தன்மை மட்டத்தில் H_0-ஐ மறுக்கிறோம்.

மாதிரி: பாட்டாளி வர்க்கத்தினர், நடுத்தரக் குடும்பத்தினர், நிலச்சுவான்தாரர்கள் என்று பலதரப்பட்ட சமூகப் பிரிவிலிருந்து வந்து பள்ளியில் படிக்கும் 413 மாணவர்கள் பொதுப்பணி, கலை, வாணிபம் என்று பல்வேறுபட்ட பாடத் திட்டங்களுக்குத் தங்கள் பெயர்களைப் பதிவு செய்துகொண்ட விவரம் பின்வருமாறு :

சமூகப் பிரிவு

பாடத் திட்டம்	பாட்டாளி வர்க்கத்தினர்	நடுத்தரக் குடும்பத்தினர்	நிலச்சுவான் தார்கள்	மொத்தம்
பொதுப்பணி	34	62	28	124
கலை	27	28	20	75
வாணிபம்	57	105	52	214
மொத்தம்	118	195	100	413

மூன்று வகைப்பட்ட பாடத்திட்டங்களுக்கு வெவ்வேறு சமூகப் பிரிவிலிருந்து வந்து தங்கள் பெயர்களைப் பதிவு செய்து கொண்ட மாணவர்களின் விகிதம் ஒரே மாதிரியாக உள்ளது என்ற சூனிய எடுகோள் H_0-ஐயும் மூன்று வகைப்பட்ட பாடத் திட்டங்களுக்கு வெவ்வேறு சமூகப் பிரிவிலிருந்து வந்து தங்கள் பெயர்களைப் பதிவு செய்துகொண்ட மாணவர்களின் விகிதம் பிரிவுக்குப் பிரிவு வேறுபட்டுள்ளது என்ற மாற்று எடுகோள் H_1-ஐயும் கருதி k சார்பற்ற கூறுகளுக்கான கைவர்க்கச் சோதனையைச் செய்வோம்.

இப்பொழுது k சார்பற்ற கூறுகளுக்குச் சோதனை செய்ய

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

என்ற சூத்திரத்தைப் பயன்படுத்தி χ^2-ன் மதிப்பைக் காண வேண்டும். கொடுக்கப்பட்டுள்ள கண்டறிந்த அலைவெண்களைக் கொண்டு எதிர்பார்க்கும் அலைவெண்களைக் கீழ்க்கண்டவாறு காண்கிறோம்.

எதிர்பார்க்கும் அலைவெண்கள்

பாடத்திட்டம்	பாட்டாளி வர்க்கத்தினர்	நடுத்தரக் குடும்பத்தினர்	நிலச்சுவான் தார்கள்
பொதுப் பணி	$\frac{118\times124}{413}=35\cdot4$	$\frac{195\times124}{413}=58\cdot5$	$\frac{100\times124}{413}=30\cdot$
கலை	$\frac{118\times75}{413}=21\cdot4$	$\frac{195\times75}{413}=35\cdot4$	$\frac{100\times75}{413}=18\cdot$
வாணிபம்	$\frac{118\times214}{413}=61\cdot1$	$\frac{195\times214}{413}=101\cdot0$	$\frac{100\times214}{413}=51\cdot$

இங்கு $k = 3, r = 3$ ஆகும்.

இப்பொழுது,

$$\chi^2 = \sum_{i=1}^{3}\sum_{j=1}^{3}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

$$= \frac{(34-35\cdot4)^2}{35\cdot4}+\frac{(62-58\cdot5)^2}{58\cdot5}+\frac{(28-30\cdot0)^2}{30\cdot0}$$

$$+\frac{(27-21\cdot4)^2}{21\cdot4}+\frac{(28-35\cdot4)^2}{35\cdot4}+\frac{(20-18\cdot2)^2}{18\cdot2}$$

$$+\frac{(57-61\cdot1)^2}{61\cdot1}+\frac{(105-101\cdot0)^2}{101\cdot0}+\frac{(52-51\cdot8)^2}{51\cdot8}$$

$$=\frac{(-1\cdot4)^2}{35\cdot4}+\frac{(3\cdot5)^2}{58\cdot5}+\frac{(-2\cdot0)^2}{30\cdot0}+\frac{(5\cdot6)^2}{21\cdot4}+\frac{(-7\cdot4)^2}{35\cdot4}$$

$$+\frac{(1\cdot8)^2}{18\cdot2}+\frac{(-4\cdot1)^2}{61\cdot1}+\frac{(4\cdot0)^2}{101\cdot0}+\frac{(0\cdot20)^2}{51\cdot8}$$

$$= \frac{1 \cdot 96}{35 \cdot 4} + \frac{12 \cdot 25}{58 \cdot 5} + \frac{4 \cdot 00}{30 \cdot 0} + \frac{31 \cdot 36}{21 \cdot 4} + \frac{54 \cdot 76}{35 \cdot 4} + \frac{3 \cdot 24}{18 \cdot 2} + \frac{16 \cdot 81}{61 \cdot 1} + \frac{16 \cdot 00}{101 \cdot 0} + \frac{0 \cdot 04}{51 \cdot 8}$$

$$= 0 \cdot 055 + 0 \cdot 209 + 0 \cdot 133 + 1 \cdot 465 + 1 \cdot 547 + 0 \cdot 178 + 0 \cdot 275 + 0 \cdot 158 + 0 \cdot 001$$

$$= 4 \cdot 021$$

அட்டவணை C-யிலிருந்து $(k-1)\ (r-1) = (3-1)\ (3-1) = 4$ வரையற்ற பாகைக்கும் $\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற மதிப்பு 9·49 ஆகும். கண்டறிந்த χ^2-ன் மதிப்பான 4·021 ஆனது அட்டவணை C-யிலுள்ள $\alpha = 0 \cdot 05$-க்கும் 4 வரையற்ற பாகைக்கும் ஏற்ற மதிப்பான 9·49ஐவிடக் குறைவாக உள்ளதால் சூனிய எடுகோளான H_0-ஐ ஏற்றுக்கொள்கிறோம். அதாவது மூன்று வகைப்பட்ட பாடத்திட்டங்களுக்கு வெவ்வேறு சமூகப் பிரிவிலிருந்து வந்து தங்கள் பெயர்களைப் பதிவு செய்து கொண்ட மாணவர்களின் விகிதத்தில் வேறுபாடு இல்லை என்ற சூனிய எடுகோளை ஏற்றுக்கொள்கிறோம்.

இடைநிலைச் சோதனையின் நீட்டிப்பு (Extension of Median Test)

இடைநிலைச் சோதனையின் நீட்டிப்பானது k சார்பற்ற சம அளவுள்ள அல்லது சம அளவற்ற கூறுகள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவையா அல்லது சமமான இடைநிலைகளைப் பெற்ற முழுமைத் தொகுதிகளிலிருந்து வரப்பெற்றவையா என்பதை அறிய உதவுகிறது. இடைநிலைச் சோதனையை k சார்பற்ற கூறுகளுக்குப் பயன்படுத்த முதலில் k கூறுகளையும் இணைத்த தொடருக்கு இடைநிலையைக் காணவேண்டும். பின்பு ஒவ்வொரு கூறிலும் உள்ள எத்தனை மதிப்புகள் இந்த இடைநிலையைவிட அதிகமாக உள்ளது என்றும் குறைவாக உள்ளது என்றும் அறிந்து $(k \times 2)$ இணைப்பட்டியலைத் தயாரிக்க வேண்டும். பிறகு k சார்பற்ற கூறுகள் இடைநிலைக்கு ஏற்றவாறு ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளைச் சோதனை செய்ய

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

என்ற சூத்திரத்தைப் பயன்படுத்தி χ^2-ன் மதிப்பைக் காண வேண்டும். இந்த χ^2 ஆனது $(k-1)\ (r-1)$ வரையற்ற பாகையைப்

பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. இந்த இடைநிலைச் சோதனையில் எப்பொழுது $r = 2$ என்பதால் $(k-1)$ $(r-1) = (k-1)$ $(2-1) = (k-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. χ^2-ன் கண்டறிந்த மதிப்பானது $(k-1)$ வரையற்ற பாகைக்கும் α என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற அட்டவணை C-யிலுள்ள மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் சூனிய எடுகோளான H_0-ஐ α என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

குறிப்பு : எதிர்பார்க்கும் அலைவெண்கள் 5-க்குக் குறைவாக இருப்பின் கைவர்க்கச் சோதனையைப் பயன்படுத்துவதால் பயனில்லை. ஆனால், அவ்வாறு சிறிய அலைவெண்களைப் பெற்ற நிரல்களை அதன் அருகிலுள்ள நிரல்களின் அலைவெண்களோடு முறையே கூட்டிய பின்பு பெறப்படும் புதிய எதிர்பார்க்கும் அலைவெண்களைப் பெற்ற இணைப்பட்டியலையும் அதற்கு ஏற்றாற் போல் கண்டறிந்த அலைவெண்களைப் பெற்ற இணைப்பட்டியலிலுள்ள முறையான நிரல்களின் அலைவெண்களை முறையே கூட்டி பெறப்பெற்ற புதிய இணைப்பட்டியலையும் பயன்படுத்தி கைவர்க்கச் சோதனை செய்யலாம்.

மாதிரி : பல்வேறு நிலைகளில் கல்வித் தகுதிகளைப் பெற்ற தாய்மார்கள் தங்கள் பிள்ளைகள் படிக்கும் பள்ளிகளுக்கு அவர்களின் கல்வி நலனுக்காக விஜயம் செய்த எண்ணிக்கை கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டது.

தாய்மார்களின் கல்வித் தகுதிகள்

தொடக்கப் பள்ளிவரை படித்தவர்கள் கூறு I	7ஆம் வகுப்புவரை படித்தவர்கள் கூறு II	10ஆம் வகுப்புவரை படித்தவர்கள் கூறு III	உயர்நிலைப் பள்ளிப் படிப்பை முடித்தவர்கள் கூறு IV	புகுமுக வகுப்புவரை படித்தவர்கள் கூறு V	பட்டப் படிப்பு முடித்த வர்கள் கூறு VI
1	1	1	3	2	6
5	2	5	2	2	2
3	1	6	9	4	
0	2	1	4	5	
2	5	7			
1	3	2			
7	0	5			
0	6	8			
3	1	0			
4	4	4			
	2	0			
		3			
		2			

பல்வேறு கல்வித் தகுதிகளைப் பெற்ற தாய்மார்கள் அவர்களின் கல்வித் தகுதிகளுக்கு ஏற்பத் தங்கள் பிள்ளைகளின் கல்வியில் அக்கறை கொண்டுள்ளனரா என்பதை இடைநிலைச் சோதனை செய்து அறிவோம். தாய்மார்கள் அவர்களின் கல்வித் தகுதி களுக்கு ஏற்பத் தங்கள் பிள்ளைகளின் நலனுக்காக அவர்களின் பள்ளிக்கு விஜயம் செய்யும் எண்ணிக்கைகளில் வேற்றுமை யில்லை என்ற சூனிய எடுகோளையும், அத்தாய்மார்களின் கல்வித் தகுதிகளுக்கு ஏற்பத் தங்கள் பிள்ளைகளின் நலனுக்காக அவர்களின் பள்ளிக்கு விஜயம் செய்யும் எண்ணிக்கைகளில் வேற்றுமை யுள்ளது என்ற மாற்று எடுகோளையும் கொள்வோம். மேலும் $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கொள்வோம். இப்பொழுது எல்லாக் கூறுகளையும் இணைத்து இடைநிலையைக் கண்டால் இடைநிலை $= 2{\cdot}5$ ஆகும். ஒவ்வொரு கூறிலும் இந்த இடைநிலை மதிப்பைவிடக் குறைவாக உள்ள மதிப்புகளின் எண்ணிக்கையையும் அதிகமாக உள்ள மதிப்புகளையும் கீழ்க்கண்ட இணைப்பட்டியலில் அமைக்கிறோம்.

தாய்மார்களின் கல்வித் தகுதிகள்

	தொடக்கப் பள்ளி வரை படித்தவர்கள் கூறு I	7 ஆம் வகுப்புவரை படித்தவர்கள கூறு II	10 ஆம் வகுப்புவரை படித்தவர்கள் கூறு III	உயர்நிலைப் பள்ளிப் படிப்பை முடித்தவர்கள் கூறு IV	புகுமுக வகுப்பு வரை படித்தவர்கள் கூறு V	பட்டப்படிப்பு முடித்தவர்கள் கூறு VI	மொத்தம்
இடைநிலை மதிப்பைவிட அதிகமாக உள்ள மதிப்புகளின் எண்ணிக்கை	5	4	7	3	2	1	22
இடைநிலை மதிப்பைவிடக் குறைவாக உள்ள மதிப்புகளின் எண்ணிக்கை	5	7	6	1	2	1	22
மொத்தம்	10	11	13	4	4	2	44

இப்பொழுது இந்தக் கண்டறிந்த அலைவெண்களுக்கேற்ற எதிர் பார்க்கும் அலைவெண்களைக் கொண்ட அட்டவணை பின்வருமாறு:

தாய்மார்களின் கல்வித் தகுதிகள்

	தொடக்கப்பள்ளி வரை படித்தவர்கள் கூறு I	7 ஆம் வகுப்புவரை படித்தவர்கள் கூறு II	10 ஆம் வகுப்புவரை படித்தவர்கள் கூறு III	உயர்நிலைப் பள்ளிப் படிப்பை முடித்தவர்கள் கூறு IV	புகுமுக வகுப்பு வரை படித்தவர்கள் கூறு V	பட்டப்படிப்பு முடித்தவர்கள் கூறு VI	மொத்தம்
இடைநிலை மதிப்பைவிட அதிகமாக உள்ள மதிப்புகளின் எண்ணிக்கை	5	5·5	6·5	2	2	1	22
இடைநிலை மதிப்பைவிடக் குறைவாக உள்ள மதிப்புகளின் எண்ணிக்கை	5	5·5	6·5	2	2	1	22
மொத்தம்	10	11	13	4	4	2	44

எதிர்பார்க்கும் அலைவெண்களைப் பெற்ற மேலுள்ள அட்டவணையில் கடைசி மூன்று நிரல்களிலுள்ள எதிர்பார்க்கும் அலைவெண்கள் 5-ஐவிடக் குறைவாக உள்ளன என்பது கவனிக்கத்தக்கது. எனவே, இந்த மூன்று நிரல்களின் முறையான அலைவெண்களைக் கூட்டி ஒரு நிரலாக அமைத்துக் கீழ்க்காணும் புதிய எதிர்பார்க்கும் அலைவெண்களையும் கண்டறிந்த அலைவெண்களையும் பெற்ற அட்டவணையை அமைக்கிறோம்.

	தொடக்கப்பள்ளி வரை படித்தவர்கள்	7-ஆம் வகுப்பு வரை படித்தவர்கள்	10-ஆம் வகுப்பு வரை படித்தவர்கள்	உயர்நிலைப் பள்ளிப் படிப்பு முடித்தவர்கள்	மொத்தம்
இடைநிலை மதிப்பை விட அதிகமாக உள்ள மதிப்புகளின் எண்ணிக்கை	5 5	5.5 4	6.5 7	5 6	22
இடைநிலை மதிப்பை விடக் குறைவாக உள்ள மதிப்புகளின் எண்ணிக்கை	5 5	5.5 7	6.5 6	5 4	22
மொத்தம்	10	11	13	10	44

இப்பொழுது,

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= \frac{(5-5)^2}{5} + \frac{(4-5\cdot5)^2}{5\cdot5} + \frac{(7-6\cdot5)^2}{6\cdot5} + \frac{(6-5)^2}{5}$$

$$+ \frac{(5-5)^2}{5} + \frac{(7-5\cdot5)^2}{5\cdot5} + \frac{(6-6\cdot5)^2}{6\cdot5} + \frac{(4-5)^2}{5}$$

$$= \frac{0}{5} + \frac{(-1\cdot5)^2}{5\cdot5} + \frac{(0\cdot5)^2}{6\cdot5} + \frac{(1\cdot0)^2}{5} + \frac{0}{5} + \frac{(1\cdot5)^2}{5\cdot5}$$

$$+ \frac{(-0\cdot5)^2}{6\cdot5} + \frac{(-1)^2}{5}$$

$$= 0 + \frac{2\cdot25}{5\cdot5} + \frac{0\cdot25}{6\cdot5} + \frac{1}{5} + 0 + \frac{2\cdot25}{5\cdot5} + \frac{0\cdot25}{6\cdot5} + \frac{1}{5}$$

$$= 0 + 0\cdot409 + 0\cdot0385 + 0\cdot2 + 0 + 0\cdot409 + 0\cdot0385 + 0\cdot2$$

$$= 1\cdot2950$$

இந்த χ^2 ஆனது $k-1 = 4-1 = 3$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாகும். இப்பொழுது அட்டவணை C-யிலிருந்து $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்திற்கும், 3 வரையற்ற பாகைக்கும் ஏற்ற மதிப்பு $7\cdot82$ ஆகும். கண்டறிந்த χ^2-ன் மதிப்பான $1\cdot295$ ஆனது அட்டவணையின் மதிப்பைவிடக் குறைவாக உள்ளதால் H_0-ஐ $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்க முடியாது என்று உய்த்துணருகிறோம்.

'க்ரஸ்கல்-வாலீஸின்' ஒருவழி பரவற்படி ஆய்வுச் சோதனை (The Kruskal-Wallis One-way Analysis of Variance Test)

கருத்தில் கொள்ளப்படும் k சார்பற்ற கூறுகள் வெவ்வேறு முழுமைத் தொகுதிகளிலிருந்து வரப்பெற்றவையா என்பதை அறிய 'க்ரஸ்கல்-வாலீஸின்' ஒருவழி பரவற்படி ஆய்வுச் சோதனை மிகவும் பயன்படுகிறது. இது தரத்தைத் (ranks) தழுவியமைந்த சோதனையாகும். இச் சோதனையின் மூலம் கூறுகளுக்கிடையேயுள்ள வேறுபாடுகள், உண்மையிலேயே முழுமைத் தொகுதியிலுள்ள

வேறுபாடுகளைக் குறிக்கின்றனவா அல்லது கூறுகளுக்கிடையே யுள்ள வேறுபாடுகள் தற்செயலாக ஏற்பட்டவையா என்பதை அறியலாம். 'க்ரஸ்கல்-வாலீஸின்' சோதனையானது k சார்பற்ற கூறுகள் ஒரே முழுமைத் தொகுதியிலிருந்து அல்லது ஒத்த முழுமைத் தொகுதிகளிலிருந்து வரப்பெற்றவை என்ற சூனிய எடுகோளைச் சோதனை செய்கிறது. ஆய்வில் கருதப்படும் மாறியானது தொடர்ச்சியான பரவலைப் பெற்றுள்ளது என்ற தற்கோளை இச் சோதனை அமைத்துக்கொள்கிறது.

கொடுக்கப்பட்ட k சார்பற்ற கூறுகளிலுள்ள மொத்தக் கண்டறிந்தவை N என்க. இந்த N கண்டறிந்தவைகளுக்குத் தரங்களைக் கொடுத்து, k கண்டறிந்த கூறுகளுக்குப் பதிலாகத் தரங்களைக் கொண்ட k கூறுகளாகக் கருதுகிறோம். மிகச்சிறிய மதிப்பிற்குத் தரம் 1 என்றும், இதைவிடச் சிறிது பெரிய மதிப் பிற்குத் தரம் 2 என்றும், தரங்களை அளிக்கிறோம். பிறகு ஒவ்வொரு கூறின் தரங்களையும் கூட்டி R_j என்று குறிக்கிறோம். இந்தத் தரங்களின் கூட்டல்களுக்குள் வேறுபாடு இருந்தால் அவை ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்ற k கூறுகளிலிருந்து வரப் பெற்றவை என்று கருத முடியுமா, முடியாதா என்பதை இச் சோதனையின் மூலம் அறியலாம்.

k சார்பற்ற கூறுகள் ஒரே முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்றால், அதாவது சூனிய எடுகோளான H_0 உண்மையெனில் 'க்ரஸ்கல்-வாலீஸ்' சோதனையில் பயன் படுத்தப்படும் H என்ற புள்ளியியல் அளவையானது $(k-1)$ வரையற்ற பாகையைப் (degrees of freedom) பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. அதாவது,

$$H = \left\{ \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} \right\} - 3(N+1)$$

ஆனது $(k-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. இதில்

k = கூறுகளின் எண்ணிக்கையாகும்.

n_j = jஆம் கூறிலுள்ள வகைகளின் (cases) எண்ணிக்கை யாகும்.

R_j = jஆம் கூறிலுள்ள தரங்களின் கூட்டலாகும்.

முக்கியமாக இச் சோதனையில் k கூறுகளின் அளவுகள் மிகச்சிறிய தாக இருத்தல் கூடாது. $n_j > 5$ எனில் அட்டவணை C-யிலுள்ள கைவர்க்க மதிப்புகளை α என்ற மிகைத்தன்மை மட்டத்திற்குப்

பெறலாம். கண்டறிந்த H-ன் மதிப்பானது அட்டவணை C-யிலுள்ள $(k-1)$ வரையற்ற பாகைக்கும் α என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் H_0-ஐ α என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். $k = 3$ என்றும் n_j-க்களின் அளவுகள் 5-ம் அதற்கும் குறைவாக இருந்தால் H-ன் மாதிரிப் பரவலுக்கு ஏற்ற தோராயமாக $(k-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவல் எனினும் மேலுள்ள சூத்திரத்திலிருந்து சரியான நிகழ்தகவு மதிப்புகள் $k = 3$-க்கும், $n_j < 5$-க்கும் அட்டவணை O-ல் கொடுக்கப்பட்டுள்ளன. இந்த அட்டவணையில் முதல் நிரலில் 5-ம் அதற்குக் குறைவான n_1, n_2, n_3 களின் ($k = 3$ என்பதால்) முடியுமான பல்வேறு அளவுகள் கொடுக்கப்பட்டுள்ளன. இரண்டாம் நிரலில் கணக்கிடப்பட்ட பல்வேறு H-ன் மதிப்புகள் கொடுக்கப்பட்டுள்ளன. மூன்றும் நிரலில் H_0-ன் கீழ் மதிப்புகள் கண்டறிந்த H-ன் மதிப்பைவிட அதிகமாக இருப்பதற்கான நிகழ்தகவு மதிப்புகள் கொடுக்கப்பட்டுள்ளன. உதாரணமாக $n_1 = 3$, $n_2 = 3$, $n_3 = 2$-க்கு $H \geqslant 6\cdot2500$ எனில் அட்டவணையிலிருந்து H_0-ஐ $0\cdot011$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கலாம்.

மாதிரி: சினிமாப் படங்களைத் தயாரிப்பதில் படத்தயாரிப்பாளர்கள், சினிமா நடிகர்களைவிடத் திறமை வாய்ந்தவர்களா என்பதை அறிய 14 சினிமா நடிகர்களைக் கொண்ட மாதிரியை நடிப்புத் தொழிலை மட்டும் விரும்பும் நடிகர்கள் என்றும், படத் தயாரிப்பாளர்களாக இருக்க விரும்பும் நடிகர்கள் என்றும், படத் தயாரிப்பாளர்கள் என்றும் மூன்று குழுக்களாகப் பிரிக்கப்பட்டு அவர்களின் திறமைகளுக்கான மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டன:

நடிப்புத் தொழிலை மட்டும் விரும்புபவர்கள்	படத் தயாரிப்பாளர்களாக இருக்க விரும்பும் நடிகர்கள்	படத் தயாரிப்பாளர்கள்
106	92	125
138	134	159
93	142	176
71	145	157
111	119	

இந்த மூன்று குழுக்களின் சராசரி மதிப்புகளுக்குள் எவ்வித வேறுபாடும் இல்லை என்ற சூனிய எடுகோளையும், மூன்று குழுக்களின் சராசரி மதிப்புகளுக்குள் வேறுபாடு உள்ளது என்ற மாற்று எடுகோளையும் கொண்டு 'க்ரஸ்கல் - வாலீஸின்' ஒருவழி பரவற்படி ஆய்வுச் சோதனையைச் செய்வோம்.

இங்கு $k = 3, n_1 = 5, n_2 = 5, n_3 = 4$ ஆகும். $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கொள்வோம். $k = 3 < 5$ என்பதால் அட்டவணை O-ஐப் பயன்படுத்த வேண்டியுள்ளது. இப்பொழுது மேலுள்ள விவரங்களுக்குத் தரங்களைக் (ranks) கொடுத்துக் கீழ்க்கண்டவாறு அட்டவணைப்படுத்துகிறோம்.

மூன்று குழுக்களுக்கும் அவர்களின் திறமைகளுக்கான தரங்கள் (Ranks)

நடிப்புத் தொழிலை மட்டும் விரும்புபவர்கள்	படத் தயாரிப்பாளர்களாக இருக்க விரும்புபவர்கள்	படத் தயாரிப்பாளர்கள்
4	2	7
9	8	13
3	10	14
1	11	12
5	6	
$R_1 = 22$	$R_2 = 37$	$R_3 = 46$

இங்கு $R_1 = 22$, $R_2 = 37$, $R_3 = 46$ ஆகியவை ஒவ்வொரு கூறுகளின் தரங்களின் கூட்டலாகும். இப்பொழுது

$$H = \left\{ \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} \right\} - 3\,(N + 1)$$

$$= \left\{ \frac{12}{14\,(14+1)} \left(\frac{(22)^2}{5} + \frac{(37)^2}{5} + \frac{(46)^2}{4} \right) - 3\,(14 + 1) \right\}$$

$$= \left\{ \frac{12}{14 \times 15} \left[\frac{484}{5} + \frac{1369}{5} + \frac{2116}{4} \right] - 3\,(15) \right\}$$

$$= \frac{2}{35}(96{\cdot}8 + 273{\cdot}8 + 529) - 45$$

$$= \frac{2}{35} \times 899{\cdot}6 - 45$$

$$= \frac{1799{\cdot}2}{35} - 45$$

$$= 51{\cdot}4 - 45$$

$= 6{\cdot}4$ ஆகும்.

அட்டவணை O-லிருந்து $n_1 = 5$, $n_2 = 5$, $n_3 = 4$-க்கு H_0-ன் கீழ் $H \geqslant 6{\cdot}4$-க்கான நிகழ்தகவு $p < 0{\cdot}049$ ஆகும். இது $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைவிடக் குறைவாக இருப்பதால் H_0-ஐ மறுக்கிறோம். அதாவது மூன்று குழுக்களின் திறமைக்கான சராசரி மதிப்புகளுக்குள் வேறுபாடு உள்ளது என்று உய்த்துணருகிறோம்.

கவனிக்கவும் : கூறுகளில் இரண்டு அல்லது அதற்கு மேற்பட்ட மதிப்புகள் ஒரே தரத்தைப் பெறும் நிலை ஏற்படுமானால் அதனை ஒத்தவைகளைப் (ties) பெற்ற கண்டறிந்த மதிப்புகள் என்று கூறுகிறோம். அப்பொழுது அம்மாதிரியான மதிப்புகளுக்குத் தரங்களைப் பகிர்ந்து கொடுக்கிறோம். எனினும் இம்மாதிரியான ஒத்தவைகளால் H ஆனது மிகவும் பாதிக்கப்படுகிறது என்பதால் திருத்தப்பட்ட H-ஐக் காண்கிறோம். இதனைக் காண்பதற்குச் சாதாரணமாகக் காண விழையும் H-க்கான சூத்திரத்தை $\left(1 - \frac{\Sigma T}{N^3 - N}\right)$ ஆல் வகுக்க வேண்டும். இங்கு $T = t^3 - t$ ஆகும். t என்பது ஒத்தவை நிகழும் எண்ணிக்கையைக் குறிக்கிறது; N என்பது k கூறுகளிலுள்ள மொத்தக் கண்டறிந்தவைகளின் எண்ணிக்கையாகும். எனவே திருத்தப்பட்ட

$$H = \frac{\left\{ \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} \right\} - 3(N+1)}{1 - \frac{\Sigma T}{N^3 - N}}$$ ஆகும்.

இம்மாதிரி திருத்தம் செய்வதால் H-ன் மதிப்பில் ஓரளவு மாற்றம் இருப்பதைக் காணலாம்.

மாதிரி : 8 விதமான உருவ அளவுகளைப் பெற்ற முயல்களின் பிறப்பு எடைகள் அவைகளின் உருவ அளவுகளால் பாதிக்கப் படுகிறதா என்பதை அறிய அவ்வுருவ அளவுகளைப் பெற்ற முயல் களின் எடைகள் பவுண்டுகளில் பின்வருமாறு குறிக்கப்பட்டன.

பல்வேறு உருவ அளவுகளைப் பெற்ற முயல் வகைகளின் எடைகள் (பவுண்டுகள்)

1	2	3	4	5	6	7	8
3·1	4·6	4·4	4·3	3·7	4·2	3·7	3·6
3·9	3·9	4·7	4·4	3·7	4·0	3·3	3·5
4·4	4·3	3·7	4·3	4·0	4·2	3·3	4·1
4·3	4·6	4·2	4·0	3·1	3·6	3·6	2·6
5·5	3·4	4·3	4·4	3·1		2·3	
4·7	3·5	4·4	3·6	3·2		2·3	
3·0	3·1	4·0	3·7				
4·4	2·7	4·5	3·9				
3·9		4·3					
2·2		4·3					

8 பல்வேறு உருவ அளவுகளைப் பெற்ற முயல்களின் சராசரிப் பிறப்பு எடைகளுக்குள் வேறுபாடு இல்லை என்ற சூனிய எடு கோளையும், அம் முயல் வகைகளின் சராசரிப் பிறப்பு எடைகள் அவைகளின் உருவ அளவுகளுக்கேற்றவாறு வேறுபடுகின்றன என்ற மாற்று எடுகோளையும் கொண்டு 'க்ரஸ்கல்-வாலீஸின்' ஒருவழி பரவற்படி ஆய்வு தரச் சோதனையைச் செய்வோம். இங்கு $N=56$, $k=8$ ஆகும். $\alpha=0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தைக் கருதுவோம். இப்பொழுது இந்த 56 மதிப்புகளுக்கும் தரங்களை

அளித்துக் கூறுகள் அமைந்த விதத்திலேயே அமைத்தால் கீழ்க் கண்ட அட்டவணையைப் பெறுகிறோம்.

பல்வேறு உருவ அளவுகளைப் பெற்ற முயல் வகைகளின் எடைகளுக்கான தரங்கள் (Ranks)

1	2	3	4	5	6	7	8
8·5	52·5	47·5	41·0	23·0	36·0	23·0	18·5
27·5	27·5	54·5	47·5	23·0	31·5	12·5	15·5
47·5	41·0	23·0	41·0	31·5	36·0	12·5	34·0
41·0	52·5	36·0	31·5	8·5	18·5	18·5	4·0
56·0	14·0	41·0	47·5	8·5		2·5	
54·5	15·5	47·5	18·5	11·0		2·5	
6·0	8·5	31·5	23·0				
47·5	5·0	51·0	27·5				
27·5		41·0					
1·0		41·0					
R_1 = 317·0	R_2 = 216·5	R_3 = 414·0	R_4 = 277·5	R_5 = 105·5	R_6 = 122·0	R_7 = 71·5	R_8 = 72·0

இங்குப் பல ஒத்தவைகளைப் (ties) பெற்ற மதிப்புகள் உள்ள தால் தரங்களைச் சமமாகப் பங்கிட்டு அளித்துள்ளோம். ஒத்தவை களுக்கான திருத்தங்களைக் கொள்ளாமல் இருந்தால்

$$H = \left\{ \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} \right\} - 3\,(N+1)$$

என்ற சூத்திரத்தின்படி

$$H = \left\{ \frac{12}{56(56+1)} \left[\frac{(317)^2}{10} + \frac{(216\cdot 5)^2}{8} + \frac{(414)^2}{10} + \frac{(277\cdot 5)^2}{8} + \frac{(105\cdot 5)^2}{6} + \frac{(122)^2}{4} + \frac{(71\cdot 5)^2}{6} + \frac{(72\cdot 0)^2}{4} \right] - 3(56+1) \right\}$$

$$= \frac{12}{56 \times 57} \left[\frac{100489}{10} + \frac{46872\cdot 25}{8} + \frac{171396}{10} + \frac{77006\cdot 25}{8} + \frac{11130\cdot 25}{6} + \frac{14884}{4} + \frac{5112\cdot 25}{6} + \frac{5184}{4} \right] - 3 \times 57 \Big\}$$

$$= \left\{\frac{12}{3192}\left[10048{\cdot}9+5859{\cdot}031+17139{\cdot}6+9625{\cdot}781 \right.\right.$$

$$\left.\left.+1855{\cdot}042+3721{\cdot}0+852{\cdot}042+1296{\cdot}0\right]-171\right\}$$

$$= \left[\frac{12}{3192}+50397{\cdot}396\right]-171$$

$$= 189{\cdot}464 - 171$$

$= 18{\cdot}464$ ஆகும்.

ஒத்தவைகளுக்கான திருத்தத்தை மேற்கொண்டால்

$$H = \frac{\left\{\frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{n_j}\right\} - 3\,(N+1)}{1 - \frac{\Sigma T}{N^3 - N}}$$

என்ற சூத்திரத்தைப் பயன்படுத்த வேண்டும். இங்கு $T = t^3 - t$ ஆகும். t என்பது ஒத்தவைகளின் எண்ணிக்கையாகும். கொடுக்கப்பட்டுள்ள அட்டவணையில்

2·3 ஆனது 2 முறையும்; அதாவது $t = 2$ என்றும்
3·1 ஆனது 4 முறையும்; அதாவது $t = 4$ என்றும்
3·3 ஆனது 2 முறையும்; அதாவது $t = 2$ என்றும்
3·5 ஆனது 2 முறையும்; அதாவது $t = 2$ என்றும்
3·6 ஆனது 4 முறையும்; அதாவது $t = 4$ என்றும்
3·7 ஆனது 5 முறையும்; அதாவது $t = 5$ என்றும்
3·9 ஆனது 4 முறையும்; அதாவது $t = 4$ என்றும்
4·0 ஆனது 4 முறையும்; அதாவது $t = 4$ என்றும்
4·2 ஆனது 3 முறையும்; அதாவது $t = 3$ என்றும்
4·3 ஆனது 7 முறையும்; அதாவது $t = 7$ என்றும்
4·4 ஆனது 6 முறையும்; அதாவது $t = 6$ என்றும்
4·6 ஆனது 2 முறையும்; அதாவது $t = 2$ என்றும்
4·7 ஆனது 2 முறையும்; அதாவது $t = 2$ என்றும்

ஏற்பட்டுள்ளன. எனவே t-களுக்கான T-க்களின் மதிப்புகளைப் பின்வருமாறு அட்டவணைப்படுத்துகிறோம்.

t	$T = t^3 - t$
2	6
4	60
2	6
2	6
4	60
5	120
4	60
4	60
3	24
7	336
6	210
2	6
2	6
	$\Sigma T = 960$

எனவே $1 - \frac{\Sigma T}{N^3 - N} = 1 - \frac{960}{(56)^3 - 56}$

$$= 1 - \frac{960}{(175616 - 56)}$$

$$= 1 - \frac{960}{175560}$$

$$= 1 - 0{\cdot}0055$$

$= 0{\cdot}9945$ ஆகும்.

இப்பொழுது ஒத்தவைகளுக்கான திருத்தத்தைச் செய்யும் H-ன் மதிப்பான

$$H = \frac{\left\{\frac{12}{N(N+1)} \sum \frac{R_j^2}{n_j}\right\} - 3(N+1)}{1 - \frac{\Sigma T}{N^3 - N}}$$

$$= \frac{18{\cdot}464}{0{\cdot}9945}$$

$= 18{\cdot}566$ ஆகும்.

இது முன்பு கண்ட திருத்தம் செய்யப்படாத மதிப்பான 18·464ஐ விடச் சிறிதளவே மாறுபட்டுள்ளது. இங்கு $k = 8$ என்பதால், மேலுள்ள H ஆனது $k-1 = 8-1 = 7$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. எனவே $\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்தில் 7 வரையற்ற பாகைகளுக்கு அட்டவணை C-யிலுள்ள மதிப்பு 14·07 ஆகும். கண்டறிந்த H-ன் திருத்தப்பட்ட மதிப்பானது, அட்டவணை C-ன் மதிப்பைவிட அதிகமாக இருப்பதால் சூனிய எடுகோளான H_0-ஐ $\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். அதாவது முயல்களின் உருவ அளவுகளுக்கேற்ப அவைகளின் எடைகள் மாறுகின்றன என்று உய்த்துணருகிறோம்.

ஜோடியாகக் கண்டறிந்தவைகளுக்குச் சார்பற்ற தன்மை பற்றிய சோதனை (Test of independence with a sample of paired observation)

ஒரே மாதிரியான பரவலைப் (identically distributed) பெற்ற k ஜோடியான கண்டறிந்தவை சார்பற்ற ஜோடிகளாக உள்ளனவா இல்லையா என்பதை அறியத் தரங்களின் அடிப் படையில் அமைந்த சோதனையை ஜெரால்டு ஜே. க்ளோஸர் (JASA, Vol. 57, 1962, pp. 116–133) என்பவர் உருவாக்கியுள்ளார்.

x_{ij} என்பது jஆம் ($j = 1, 2$) கண்டறிந்த மதிப்பை iஆம் ($i = 1, 2, \ldots. k$) ஜோடியைக் குறிக்கட்டும். இதில் கண்டறிந் தவை எந்த வரிசையில் தேர்ந்தெடுக்கப்பட்டுள்ளன என்பதை i, j ஆகியவை குறிக்கின்றன. எனவே, ஆய்வில் கொள்ளப்படும் மாதிரியானது,

$$(x_{11}, x_{12}, x_{21}, x_{22}, \ldots\ldots x_{k1}, x_{k2}) \qquad (1)$$

என்று அமைந்திருக்கும்.

பௌதிக அளவுகளையோ அல்லது பண்பளவுகளையோ பெற்ற கூறுகளைக் கருதும்பொழுது, ஜோடியான கண்டறிந்தவைகளைப் பெறவேண்டிய நிலை ஏற்படுகிறது. உதாரணமாக இரு குவிவ லென்சுகளை (biconcave lenses) உற்பத்தி செய்து விற்பனைக்கு அனுப்புதற்கு முன்பு அந்த லென்சுகளின் இருபுறமும் உள்ள குறைகளின் எண்ணிக்கையைக் குறிக்கும்பொழுது (x_{i1}, x_{i2}) என்ற ஜோடியான கண்டறிந்தவைகளைப் பெற்ற மாதிரியைப் பெறுகிறோம். மற்றோர் உதாரணமாக, இரட்டைப் பிறவி களாகப் பிறந்த குழந்தைகளின் இரத்த அழுத்தத்தைப் பதிவு செய்து இந்த இரட்டைக் குழந்தைகள் இரத்த அழுத்தத்திற்குச்

சார்பற்ற தன்மையைப் பெற்றுள்ளனவா என்பதை அறிய விழையும்போது ஜோடியான கண்டறிந்தவைகளைப் பெறுகிறோம். இதனை அறிய, கண்டறிந்த ஜோடிகளைத் தரங்களின் அடிப் படையில் கருதுகிறோம்.

'க்ரஸ்கல்-வாலீஸின்' சோதனையில் பயன்படுத்தப்பட்ட H என்ற புள்ளியியல் அளவையின் சிறப்பான ஒரு புள்ளியியல் அளவையை இங்குப் பயன்படுத்திச் சோதனை செய்கிறோம்.

ஜோடியாக அமைந்த கண்டறிந்தவை சார்பற்றன என்ற சூனிய எடுகோளை, அந்தக் கண்டறிந்த ஜோடிகள் ஒருபடித் தானவை (homogenous) என்ற ஒரு முனைக்கான மாற்று எடு கோளுக்கு எதிராகக் கொண்டு இச் சோதனையை மேற்கொள் கிறோம்.

இப்பொழுது கண்டறிந்த k ஜோடிகளின் மதிப்புகளுக்குத் தரங்களை இறங்கு வரிசையிலோ (அல்லது ஏறு வரிசையிலோ) அமைத்து அவைகளுக்கு $(1, 2, \ldots, n)$ என்று குறிக்கிறோம். இதில் $n = 2k$ ஆகும். தற்பொழுது கண்டறிந்தவைகளில் ஒத்தவை (ties) இல்லை என்று கொள்வோம். இப்பொழுது r_{ij} என்பது iஆம் ஜோடியின் jஆம் கண்டறிந்த மதிப்புகளுக்கான k ஜோடி களுக்கான தரங்கள் எனில், அவைகளை

$$(r_{11}, r_{12}, r_{21}, r_{22}, \ldots\ldots\ldots r_{k1}, r_{k2}) \qquad (2)$$

என்று பெறுகிறோம். இப்பொழுது D என்ற புள்ளியியல் அவையை

$$D = \Sigma\,(r_{i1} - r_{i2})^2 \qquad (3)$$

$$(i = 1, 2, \ldots\ldots, k)$$

என்று குறிக்கிறோம்.

D-ன் மிகக் குறைந்த மதிப்பானது ஜோடியான கண்டறிந்தவை சார்பற்றன என்று எடுகோளை மறுக்க வேண்டும் என்பதை இச்சோதனைக்கு உணர்த்துகிறது.

D-ன் மதிப்பைவிட அதிகமாக இருக்கும் D^* என்ற தீர்வுகட்ட மான மதிப்புகள் $\alpha = 0{\cdot}005,\ 0{\cdot}01,\ 0{\cdot}025,\ 0{\cdot}05,\ 0{\cdot}10$ என்ற மிகைத்தன்மை மட்டங்களுக்கு, கீழுள்ள அட்டவணையில் $k = 3$ லிருந்து $k = 15$ வரை கொடுக்கப்பட்டுள்ளன.

D-ன் தோராயமான தீர்வு கட்டமான மதிப்பு *D**

k	$\alpha=0{\cdot}005$	$\alpha=0{\cdot}01$	$\alpha=0{\cdot}025$	$\alpha=0{\cdot}05$	$\alpha=0{\cdot}10$
3	—	—	—	9	9
4	—	10	12	16	20
5	13	17	23	31	41
6	26	34	46	58	76
7	49	61	81	101	129
8	84	102	132	162	202
9	135	159	203	243	297
10	202	236	294	350	420
11	289	337	411	483	573
12	402	462	558	650	762
13	541	617	737	849	987
14	710	802	950	1086	1254
15	915	1027	1205	1369	1569
*Z**	2·58	2·33	1·96	1·64	1·28

சூத்திரம் (3) லிருந்து கண்ட *D*-ன் மதிப்பானது *D** என்ற தீர்வு கட்டமான மதிப்பைவிடக் குறைவாக இருந்தால் ஜோடியான கண்டறிந்தவை சார்பற்றன என்ற சூனிய எடுகோளை α என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். மற்றபடி சூனிய எடு கோளை ஏற்றுக்கொள்கிறோம்.

$k > 15$ எனில்,

$$Z = \frac{k^2\,(2k+1) - 3D}{k\,(2\,k+1)\sqrt{k-1}} \qquad (4)$$

என்ற புள்ளியியல் சோதனை அளவையானது *N* (0, 1) என்ற தோராயமான இயல்நிலைப் பரவலாக அமைகிறது. கண்டறிந்த *Z*-ன் மதிப்பானது மேலே கொடுக்கப்பட்டுள்ளது. அட்டவணையி லுள்ள *Z**-ன் மதிப்பைவிட அதிகமாக இருந்தால் ஒருமுனைச் சோதனைக்கு ஜோடியான கண்டறிந்தவை சார்பற்றன என்ற சூனிய எடுகோளை மறுக்கிறோம். மற்றபடி சூனிய எடுகோளை α என்ற மிகைத்தன்மை மட்டத்தில் ஏற்றுக்கொள்கிறோம்.

மாதிரி : மரப்பலகைகளின் இரு ஓரங்களில் பொருத்தப் பட்டுள்ள அலுமினியத் தகடுகளின் நீளங்கள் (அங்குலத்தில்)

ஜோடியாக 12 மரப்பலகைகளுக்குப் பின்வருமாறு குறிக்கப் பட்டன.

(0·538, 0·545)	(0·552, 0·556)	(0·544, 0·542)
(0·547, 0·566)	(0·563, 0·568)	(0·570, 0·567)
(0·579, 0·564)	(0·585, 0·559)	(0·574, 0·554)
(0·549, 0·580)	(0·558, 0·561)	(0·571, 0·562)

மேலுள்ள ஜோடியான கண்டறிந்த மதிப்புகள் புள்ளியியல் சார்பற்ற தன்மையைப் பெற்று அமைகின்றனவா என்பதை இச் சோதனையைப் பயன்படுத்திக் காண்போம். இப்பொழுது இந்த 24 மதிப்புகளையும் இறங்கு வரிசையில் தரப்படுத்தினல்,

(24, 21)	(18, 16)	(22, 23)
(20, 9)	(11, 17)	(6, 8)
(3, 10)	(1, 14)	(4, 17)
(19, 2)	(15, 13)	(5, 12)

என்று பெறுகிறோம்.

எனவே, $D = (24-21)^2 + (20-9)^2 + \dots\dots\dots\dots + (4-17)^2 + (5-12)^2$

$= 884$ ஆகும்.

மிகைத்தன்மை மட்டத்தை $\alpha = 0{\cdot}05$ எனக்கொண்டால், $k = 12$-க்கு அட்டவணையின் மதிப்பு $D^* = 650$ ஆகும். கண்டறிந்த D-ன் மதிப்பான 884 ஆனது $D^* = 650$ என்ற அட்டவணை மதிப்பைவிட அதிகமாக உள்ளதால் ஜோடியான கண்டறிந்தவை சார்பற்றன என்ற சூனிய எடுகோளை $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் ஏற்றுக்கொள்கிறோம்.

குறிப்பு : கண்டறிந்த ஜோடிகளிலுள்ள மதிப்புகள் ஒத்தவைகளாக (tied) இருந்தால், அந்த ஜோடிகளின் மதிப்புகளுக்குத் தரங்களைச் சமமாகப் பங்கிட்டு அளிக்கிறோம். உதாரணமாக, L குழுக்களில் $t_1, t_2, \dots\dots, t_L$ என்று ஒத்தவை இருந்தால் பின்பு,

$$T = 1 - \left[\frac{\Sigma\,(t_i^3 - t_i)}{2k\,(2k+1)\,(2k-1)} \right]; \; (i = 1, 2, \dots\dots, L)$$

என்று திருத்தத்திற்கான புள்ளியியல் அளவையைக் காண்கிறோம்.

பிறகு,

$$Z = \frac{k^2(2k+1)T - 3D}{k(2k+1)T\sqrt{k-1}} \tag{5}$$

இதில் $T=1$ எனில். சூத்திரம் (5), சூத்திரம் (4)-க்குச் சமமாகிறது.

மாதிரி : 19 ஜோடிகளின் தனித்தனியான இரட்டைக் குழந்தைகளின் நரம்புவழிநோய்ப் பண்புகளின் (Neurotic traits) மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டுள்ளன.

(17, 18)	(17, 17)	(30, 30)	(5, 2)
(18, 26)	(14, 12)	(26, 20)	(15, 26)
(13, 15)	(11, 19)	(24, 19)	(5, 11)
(3, 8)	(12, 8)	(37, 15)	(13, 15)
(18, 15)	(14, 11)	(11, 7)	

மேலுள்ள விவரங்களிலிருந்து ஜோடிகளின் கண்டறிந்த மதிப்புகள் சார்பற்றனவா என்பதை அறிவோம்.

இப்பொழுது கொடுக்கப்பட்டுள்ள கண்டறிந்தவைகளைத் தரப்படுத்தி அமைத்தால்,

(24, 27)	(24, 24)	($36\frac{1}{2}$, $36\frac{1}{2}$)	($3\frac{1}{2}$, 1)
(27, 34)	($16\frac{1}{2}$, $12\frac{1}{2}$)	(34, 31)	(20, 34)
($14\frac{1}{2}$, 20)	($9\frac{1}{2}$, $29\frac{1}{2}$)	(32, $29\frac{1}{2}$)	($3\frac{1}{2}$, $9\frac{1}{2}$)
(2, $6\frac{1}{2}$)	($12\frac{1}{2}$, $6\frac{1}{2}$)	(38, 20)	($14\frac{1}{2}$, 20)
(27, 20)	($16\frac{1}{2}$, $9\frac{1}{2}$)	($9\frac{1}{2}$, 5)	

என்று பெறுகிறோம்.

எனவே, $D = (24-27)^2 + (27-34)^2 + (14\frac{1}{2}-20)^2$
$+ (2-6\frac{1}{2})^2 + (27-20)^2 + (24-24)^2$
$+ (16\frac{1}{2}-12\frac{1}{2})^2 + \ldots\ldots\ldots\ldots$
$+ (3\frac{1}{2}-9\frac{1}{2})^2 + (14\frac{1}{2}-20)^2$

$= 1286 \cdot 50$ ஆகும்.

கொடுக்கப்பட்ட ஜோடியான மதிப்புகளில் 3, 5, 7, 8, 12, 13, 19 ஆகிய மதிப்புகள் இருமுறை நிகழ்ந்துள்ளன. எனவே,

இரு ஒத்தவை 7 முறை ஏற்பட்டுள்ளன. 17, 18, 21 ஆகிய மதிப்புகள் மூன்று முறை நிகழ்ந்துள்ளன. எனவே, மூன்று ஒத்தவை 3 முறை ஏற்பட்டுள்ளன. மற்றபடி 11 என்ற மதிப்பு 4 முறை நிகழ்ந்துள்ளது. எனவே, நான்கு ஒத்தவை ஒரு முறை ஏற்பட்டுள்ளது. இறுதியாக 15 ஆனது 5 முறை ஏற்பட்டுள்ளது. எனவே, ஐந்து ஒத்தவை 1 முறை ஏற்பட்டுள்ளது.

இப்பொழுது,

$$\Sigma(t_i^3 - t_i) = 7(2^3 - 2) + 3(3^3 - 3) + 1(4^3 - 4) + 1(5^3 - 5)$$

$$= 7(6) + 3(24) + (60) + 120$$

$$= 294 \text{ ஆகும்.}$$

எனவே,

$$T = 1 - \left[\frac{\Sigma(t_i^3 - t_i)}{2k(2k+1)(2k-1)}\right], \quad (i = 1, 2, \ldots L)$$

$$= 1 - \left\{\frac{294}{2(19)[2(19)+1][2(19)-1]}\right\}$$

$$= 0 \cdot 99464 \text{ ஆகும்.}$$

எனவே,

$$Z = \frac{k^2(2k+1)T - 3D}{k(2k+1)T\sqrt{k-T}}$$

$$= \frac{(19)^2(39)(0 \cdot 99464) - 3(1286 \cdot 5)}{19(39)(0 \cdot 99464)\sqrt{18}}$$

$$= 3 \cdot 24 \text{ ஆகும்.}$$

D-ன் தோராயமான தீர்வு கட்டமான மதிப்பு அட்டவணையிலிருந்து $\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்திற்கான $Z^* = 1 \cdot 64$ ஆகும். கண்டறிந்த Z-ன் மதிப்பான $3 \cdot 24$ ஆனது, அட்டவணை மதிப்பான $Z^* = 1 \cdot 64$ஐ விட அதிகமாக உள்ளதால் ஒருமுனைச் சோதனைக்கு, ஜோடியான கண்டறிந்தவை சார்பற்றன என்ற சூனிய எடுகோளை $\alpha = 0 \cdot 05$ என்ற மிகைத் தன்மை மட்டத்தில் மறுக்கிறோம்.

D-க்கும், H-க்கும் இடையேயுள்ள தொடர்பு

'க்ரஸ்கல்-வாலீஸ்' சோதனையில் பயன்படுத்தப்படும் புள்ளியியல் அளவையான H ஆனது மேற்சொன்ன சோதனையின் புள்ளியியல் அளவையான D-யோடு பின்வருமாறு தொடர்பு கொண்டுள்ளது.

$$H = \frac{k(4k^2 - 1) - 3D}{k(2k + 1)}$$

அல்லது,

$D = \frac{1}{3}[k(4k^2 - 1) - k(2k + 1)H]$ ஆகும்.

பயிற்சி 5

1. *k* சார்பற்ற கூறுகளுக்கான கைவர்க்கச் சோதனையின் அமைப்பையும், அச் சோதனையின் பயன்முறைகளையும் விளக்குக.

2. வெவ்வேறு நிலைகளில் கணிதத்திறனும், புள்ளியியலிலுள்ள ஆர்வமும் பெற்ற நபர்களின் விவரங்கள் (3×3) இணைப்புப் பட்டியலில் கொடுக்கப்பட்டுள்ளன.

		கணிதத் திறமை		
		குறைந்த	சுமாரான	அதிகமான
புள்ளியியலிலுள்ள ஆர்வம்	குறைந்த	63	42	15
	சுமாரான	58	61	31
	அதிகமான	14	47	29

நபர்களின் கணிதத்திறமைக்கும், புள்ளியியலிலுள்ள ஆர்வத்திற்குமிடையே தொடர்பு ஏதேனும் உண்டா இல்லையா என்பதை *k* கூறுகளுக்கான கைவர்க்கச் சோதனையின் மூலம் $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் ஆய்க.

3. ஒரு நாட்டின் பல்வேறு பகுதிகளில் வாழும் மக்கள் 'நிர்ணயிக்கப்படாத மருந்து' (Non - prescribed medicine) வகைகளை வாங்கும் விவரங்கள், பின்வரும் அட்டவணையில் கொடுக்கப்பட்டுள்ளன.

	வடக்கு	கிழக்கு	தெற்கு	மேற்கு
மருந்துக் கடைகள்	219	200	181	180
மளிகைக் கடைகள்	39	52	89	60
மற்றவை	42	48	30	60

நாட்டின் வெவ்வேறு பகுதிகளிலிருந்து 'நிர்ணயிக்கப்படாத' மருந்து வகைகளை வாங்குவதில் மக்கள் வேறுபடுகின்றனரா என்பதை $\alpha = 0\cdot05$ என்ற மிகைத்தன்மை மட்டத்தில் ஆய்க.

4. தமிழ்நாட்டின் ஒரு பகுதியில் மூன்று வகைப்பட்ட பரம்பரை நபர்கள், அவர்களின் இரத்தப் பிரிவுகளுக்கு ஏற்பக் கீழ்க்கண்டவாறு அட்டவணைப்படுத்தப்பட்டனர்.

	இரத்தப் பிரிவுகள்			
	O	A	B	AB
பரம்பரை 1	176	148	96	72
பரம்பரை 2	78	50	45	12
பரம்பரை 3	15	19	8	7

பரம்பரைகளுக்கும், இரத்தப் பிரிவுகளுக்குமிடையே தொடர்பு இருக்கிறதா, இல்லையா என்பதை k கூறுகளுக்கான கைவர்க்கச் சோதனையைப் பயன்படுத்தி $\alpha = 0\cdot01$ என்ற மிகைத் தன்மை மட்டத்தில் ஆய்க.

5. A, B, C என்ற மூன்று சலவை சோப்புக்கட்டிகள், அழுக் கடைந்த வெள்ளைத் துணிகளை வெண்மையாக்கும் தன்மைகுறித் துக் குளிர்ந்த நீர், வெதுவெதுப்பான நீர், சுடுநீர் ஆகியவைகளில் சோதித்துப் பார்த்ததில் கீழ்க்கண்ட விவரங்கள் கிடைத்தன.

	சோப்பு A	சோப்பு B	சோப்பு C
குளிர்ந்த நீர்	45	43	55
வெதுவெதுப்பான நீர்	37	40	56
சுடுநீர்	42	44	46

துணிகளை வெண்மையாக்கும் தன்மையில் சோப்பு வகைகள், நீரின் வெவ்வேறு வகைகளைச் சார்ந்துள்ளனவா இல்லையா என்பதை கைவர்க்கச் சோதனைமூலம் ஆய்க.

6. *k* சார்பற்ற கூறுகளுக்கான இடைநிலைச் சோதனையைப் பயன்படுத்தும் முறையை விளக்குக.

7. பல்வேறு நிலைகளில் கல்வித் தகுதிகளைப் பெற்ற பெற்றோர்கள், தங்கள் பிள்ளைகளின் நலனைக் கருதி, தலைமை ஆசிரியருடன் கலந்துரையாடிய சந்தர்ப்பங்களின் எண்ணிக்கை பின்வருமாறு :

பெற்றோர்களின் கல்வித் தகுதிகள்

ஆரம்பப் பள்ளி வரை படித்தவர்கள்	8ஆம் வகுப்பு வரை படித்தவர்கள்	உயர்நிலைப் பள்ளி இறுதி வகுப்பு வரை படித்தவர்கள்	பட்டப்படிப்பு படித்தவர்கள்	மேற்பட்டப் படிப்பு படித்தவர்கள்
4	2	4	6	9
6	6	5	9	9
3	9	3	4	8
2	2	2	8	9
0	4	1	9	8
2	5	4	9	9
1	3	5	8	10
6	2	6	5	8
0	0	8	7	
4	0	9		
5	5	2		
		1		
		1		
		5		

பல்வேறு கல்வித் தகுதிகளைப் பெற்ற பெற்றோர்களின் தகுதிகளுக்கேற்ப, அவர்கள் தங்கள் பிள்ளைகளின் கல்வியில் அக்கறை கொண்டுள்ளனரா என்பதை இடைநிலைச் சோதனையைப் பயன்படுத்தி ஆய்க.

8. 'க்ரஸ்கல்-வாலீஸின்' ஒருவழி பரவற்படி ஆய்வுச் சோதனையைப்பற்றி விளக்குக.

9. திரைப்படப் பாடல்களுக்கு இசை அமைப்பதில் இசையமைப்பாளர்கள், பாடகர்களைவிடத் திறமை வாய்ந்தவர்களா என்பதை அறிய 18 பாடகர்களைக் கொண்ட கூறு, பாடும் தொழிலைமட்டும் விரும்புபவர்கள் என்றும், இசையமைப்பாளர்களாக ஆக விரும்பும் பாடகர்கள் என்றும், இசையமைப்பாளர்கள் என்றும் மூன்று குழுக்களாகப் பிரிக்கப்பட்டு, அவர்களின் திறமைகளுக்கான மதிப்புகள் கீழ்க்கண்டவாறு அட்டவணைப் படுத்தப்பட்டன.

பாடும் தொழிலை மட்டும் விரும்புபவர்கள்	இசையமைப்பாளராக இருக்க விரும்புபவர்கள்	இசையமைப்பாளர்கள்
112	99	159
124	68	175
98	110	148
46	85	88
121	114	72
92	146	
	56	

இம் மூன்று குழுக்களின் சராசரி மதிப்புகளுக்குள் வேறுபாடு உள்ளதா இல்லையா என்பதை 'க்ரஸ்கல்-வாலீஸ்' ஒருவழி பரவற்படி ஆய்வுச் சோதனையைப் பயன்படுத்தி ஆய்க.

10. 6 விதமான உருவ அளவுகளைப் பெற்ற செம்மறி ஆடுகளின் பிறப்பு எடைகள் (கிலோ கிராமில்) அவைகளின் உருவ அளவுகளுக்கேற்பக் கீழ்க்கண்டவாறு குறிக்கப்பட்டுள்ளன.

பல்வேறு உருவ அளவுகளைப்பெற்ற செம்மறி ஆட்டு வகைகளின் எடைகள் (கிலோ கிராமில்)

(1)	(2)	(3)	(4)	(5)	(6)
9·1	10·1	10·5	10·9	7·6	7·9
9·8	9·6	10·7	9·7	9·8	8·8
10·3	8·8	11·0	9·9	10·2	10·4
10·2	9·2	10·2	8·8	11·1	11·0
10·6	7·5	9·8	10·7		9·6
8·6	6·8	8·9			9·2
7·9		7·5			
8·9					

பிறப்பு எடைகளுக்கும், உருவ அளவுகளுக்குமிடையே வேறுபாடு உள்ளதா இல்லையா என்பதை 'க்ரஸ்கல்-வாலீஸ்' ஒருவழிப் பரவற்படி ஆய்வு தரச் சோதனையைப் பயன்படுத்தி ஆய்க.

11. ஜோடியாகக் கண்டறிந்த விவரங்களுக்குச் சார்பற்ற தன்மையை அறிய உதவும் சோதனையைப்பற்றி விவரிக்கவும். இச் சோதனையின் புள்ளியியல் அளவைக்கும், 'க்ரஸ்கல்-வாலீஸ்' சோதனையின் புள்ளியியல் அளவைக்கும் உள்ள தொடர்பு என்னவென்பதைக் கூறு.

12. குவிவ லென்சுகளை உற்பத்தி செய்ததில் 14 குவிவ லென்சுகளின் இருபுறமும் உள்ள குறைகளின் எண்ணிக்கைகள் கீழ்க்கண்டவாறு குறிக்கப்பட்டன.

(23,2),	(5,24),	(21,7),	(12,10)
(16,9),	(8,11),	(13,14),	(17,26)
(6,18),	(17,19),	(1,20),	(3,21)
(22,4),	(25,15)		

மேலுள்ள விவரங்களைக் கொண்டு ஜோடியாகக் கண்டறிந்த லென்சுகளின் குறைகளின் எண்ணிக்கைகளுக்குள் சார்பற்ற தன்மைபற்றி ஆய்க.

6. ஒட்டுறவின் அளவைகளும் அவைகளுக்கான மிகைத்தன்மைச் சோதனைகளும்

(Measures of Correlation and their Tests of Significance)

ஒரு கூறின் மதிப்புகளுக்கு இணையாக மற்றொரு கூறின் மதிப்புகள் வேறுபட்டு அமைந்தால் அவ்விரு கூறுகளைத் தொடர்புடைய கூறுகள் என்கிறோம். கூறுகளுக்கிடையேயுள்ள தொடர்பை அல்லது உறவை 'ஒட்டுறவு' (Correlation) என்கிறோம். கொடுக்கப்படும் இரு கூறுகளின் மதிப்புகளுக்கிடையேயுள்ள உறவின் அளவான ஒட்டுறவுக் கெழுவைப் (Correlation coefficient) பயன்படுத்திப் பல்வேறு ஆய்வுகளை மேற்கொள்கிறோம். இவ்வத்தியாயத்தில் சுட்டுறுப்பைச் சாராத ஒட்டுறவின் அளவைகளுக்கான புள்ளியியல் சோதனைகளைப்பற்றிக் காண்போம்.

இரு கூறுகளின் மதிப்புகளுக்கிடையேயுள்ள தொடர்பின் அளவைப்பற்றி அறிவதும் ஒரு முழுமைத் தொகுதியில் அக் கூறுகளின் மதிப்புகளுக்கிடையே தொடர்பு இருக்கிறதா என்று சோதனை செய்து அறிவதும் வெவ்வேறுவகைப்பட்டதாகும். சாதாரணமாக இரு கூறுகளின் மதிப்புகளுக்கிடையேயுள்ள ஒட்டுறவுக்கெழு பூச்சியமல்லாமல் -1 விருந்து $+1$ வரை இருந்தால் அக் கூறுகளின் மதிப்புகளுக்கிடையே தொடர்பு இருக்கிறது என்று கொள்கிறோம். ஆனால், மாதிரி மதிப்புகளுக்கிடையேயுள்ள கண்டறிந்த தொடர்பானது (observed association) தற்போது ஆய்வில் கொள்ளப்பட்ட மாறிகள், பெரும்பாலும் முழுமைத் தொகுதியில் (population) தொடர்புடையனவாக அமைந்துள்ளன என்று உணர்த்துகிறதா என்பதை அறிவது பயனுள்ளதொன்றாகும். ஒட்டுறவுக்கெழுவே தொடர்பின் அளவைத் (degree of association) தெரிவிக்கிறது. கண்டறிந்த ஒட்டுறவுக் கெழுவிற்கு மிகைத்தன்மைச் சோதனையை α என்ற மிகைத்தன்மை மட்டத்தில் மேற்கொள்ளும்போது, கூறுகள் தேர்ந்தெடுக்கப்பட்ட முழுமைத்

தொகுதியில் மாறிகளுக்கிடையே தொடர்புள்ளதா என்பதை அறியலாம்.

சுட்டுறுப்பைச் சார்ந்த முறைகளில் சாதாரண ஒட்டுறவான r என்ற பியர்ஸன் ஒட்டுறவுக் கெழுவைப் பயன்படுத்துகிறோம். இந்தப் புள்ளியியல் அளவை (test statistic) சமமான இடைவெளி அளவைப் (equal interval scale) பெற்றிருக்க வேண்டும். இந்தக் கண்டறிந்த r-க்கு மிகைத்தன்மைச் சோதனை செய்யும்போது கூறுகளின் மதிப்புகள் இருமாறி இயல்நிலைப் பரவலைப் (bivariate normal distribution) பெற்ற முழுமைத் தொகுதியிலிருந்து வரப்பெற்றவை என்ற தற்கோளை மேற்கொள்கிறோம். கொடுக்கப் பட்டுள்ள விவரத்திற்கு இயல்நிலைத் தற்கோள் உண்மைக்குப் புறம் பானதாக இருக்கும் தருணத்தில் சுட்டுறுப்பைச் சாராத ஒட்டுறவுக் கெழுக்களில் ஒன்றைப் பயன்படுத்தி அதற்கேற்ற மிகைத்தன்மைச் சோதனையை மேற்கொள்ளலாம். சுட்டுறுப்பைச் சாராத ஒட்டுறவு அளவைகள் பெயரளவான அளவுத் திட்டத் திற்கும் (Nominal Scale) வரிசை அல்லது தர அளவுத் திட்டத் திற்கும் (Ordinal or Ranking Scale) அமைந்துள்ளன. சுட்டுறுப் பைச் சாராத ஒட்டுறவு அளவைகளுக்கான சோதனைகள் அனைத்தும் தற்கோள்கள் எதனையும் மேற்கொள்ளுவதில்லை. மாதிரிகளின் அளவுகள் மிகச் சிறியனவாக இருக்கும்போது சுட் டுறுப்பைச் சாராத ஒட்டுறவு அளவுகளைக் கணக்கிடுவதும் மிகைத்தன்மைச் சோதனையை மேற்கொள்வதும் மிக எளிதாக உள்ளன.

இணைப்புக் கெழு C (The Contingency Coefficient C)

பண்புகளைப் (attributes) பெற்ற இருமாறிகளுக் (variables) கிடையேயுள்ள தொடர்பின் அளவை இணைப்புக் கெழு C என்கிறோம். பண்புகளைப் பற்றிய தகவல் வரிசைப்படுத்தப் படாத தொடராக உள்ள அலைவெண்களைக் கொண்டிருந்தால் இந்த இணைப்புக் கெழுவைப் பயன்படுத்தலாம். இரு மாறிகளுக் கிடையேயுள்ள இணைப்புக் கெழுவைக் காண A எனும் ஒரு மாறியை அதன் பண்பின் அடிப்படையில் $A_1, A_2, \ldots\ldots, A_k$ என்று k பிரிவுகளிலும், B எனும் மற்றொரு மாறியை அதன் பண்பின் அடிப்படையில் $B_1, B_2, \ldots\ldots, B_r$ என்று r பிரிவுகளிலும் அமைப்பதாகக் கொள்வோம். இப்பொழுது இருவழிப் பட்டியல் ஒன்றில் $A_1, A_2, \ldots\ldots, A_k$ பிரிவுகளை நிரைகளிலும், $B_1, B_2, \ldots\ldots, B_r$ பிரிவுகளை நிரல்களிலும் அமைக்கும்போது $(k \times r)$ அறைகளைப் (cells) பெறுகிறோம்.

இணைப்புப் பட்டியல்

	A_1	A_2	....	A_k	மொத்தம்
B_1	$(A_1 B_1)$	$(A_2 B_1)$	....	$(A_k B_1)$	
B_2	$(A_1 B_2)$	$(A_2 B_2)$	....	$(A_k B_2)$	
⋮	⋮	⋮		⋮	
⋮	⋮	⋮		⋮	
B_r	$(A_1 B_r)$	$(A_2 B_r)$	....	$(A_k B_r)$	
மொத்தம்					N

ஒரு கூறில் இடம் பெறும் N உறுப்புகள் ஒவ்வொன்றையும் A, B மாறிகளின் அடிப்படையில் அவைகளுக்குரிய அறையில் வகைப் படுத்துவோம். A_i நிரையிலும், B_j நிரலிலும் உள்ள அறையை (A_i, B_j) அறை எனக் கொண்டு (A_i, B_j) அறையில் இடம் பெறும் கண்டறிந்த (observed) உறுப்புகளின் எண்ணிக்கை O_{ij} எனக் குறிப்பிடுவோம்.

இவ் அட்டவணையில் இரு மாறிகளுக்குமிடையே தொடர்பு இல்லை என்ற தற்கோளின் அடிப்படையில் ஒவ்வோர் அறையிலும் E_{ij} என்ற அலைவெண்களை எதிர்பார்க்கிறோம். O_{ij} என்ற கண்டறிந்த அலைவெண்ணுக்கும் E_{ij} என்ற எதிர்பார்க்கும் அலை வெண்ணுக்குமிடையே அதிகமான வித்தியாசங்களிருந்தால் இரு மாறிகளுக்குமிடையேயுள்ள தொடர்பின் அளவு அதிகமாக இருக்கும். எனவே, இணைப்புக்கெழுவான C-ன் மதிப்புப் பெரிய தாக இருக்கும்.

இரு கூறுகளின் பண்புகளுக்கிடையேயுள்ள தொடர்பு அளவையை இணைப்புப் பட்டியலிலிருந்து காணலாம். இதைக் காண

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

என்ற சூத்திரத்தைப் பயன்படுத்துகிறோம். இதில்

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ ஆகும்.}$$

இந்த χ^2 என்ற புள்ளியியல் அளவையைக் கணக்கிடுவது பற்றி முன்பே கண்டோம். $(r - 1)$ $(k - 1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது.

இணைப்புக் கெழுவிற்கு மிகைத்தன்மைச் சோதனை (Testing the Significance of the Contingency Coefficient)

கொடுக்கப்பட்ட கண்டறிந்த அலைவெண்களை $(k \times r)$ இணைப் புப் பட்டியலில் அமைக்கிறோம். இதில் k என்பது, அட்டவணையி லுள்ள மொத்த நிரைகளையும், r என்பது அவ்வட்டவணையி லுள்ள மொத்த நிரல்களையும் குறிக்கின்றன. பிறகு H_0-ன் கீழ், எதிர்பார்க்கும் அலைவெண்களைக் கண்ட பிறகு χ^2-ன் மதிப்பைக் காணவேண்டும். இந்த χ^2-ன் மதிப்பைப் பயன்படுத்தி இணைப்புக் கெழுவான C-ஐக் காணவேண்டும்.

கண்டறிந்த C-ன் மதிப்பானது முழுமைத் தொகுதியிலிருந்து வருவிக்கப்பட்ட மாதிரிகளான இரு மாறிகளுக்கிடையே தொடர்பு உள்ளதா என்பதைச் சோதனை செய்ய $(k—1)$ $(r—1)$ வரையற்ற பாகையைப் பெற்ற கண்டறிந்த χ^2-ன் மதிப்பைவிட அதிகமாக இருப்பதற்கு H_0-ன்கீழ் தொடர்புடைய நிகழ்தகவின் மதிப்பை, அட்டவணை C-லிருந்து பெறவேண்டும். காணும் நிகழ்தகவின் மதிப்பு α என்ற மிகைத்தன்மை மட்டத்திற்குச் சமமாகவோ அல்லது குறைவாகவோ இருந்தால் H_0-ஐ மறுக் கிறோம்.

மாதிரி : 200 மாணவர்கள் கொண்ட ஒரு குழுவில், பொது அறிவுக்கும், சமூக நடத்தைக்குமான புள்ளி விவரங்கள் பின்வரும் இணைப்புப் பட்டியலில் கொடுக்கப்பட்டுள்ளன.

		சமூக நடத்தை			மொத்தம்
		நல்ல	சுமாரான	கெட்ட	
பொது அறிவு	நல்ல	40	30	20	90
	சுமாரான	10	10	10	30
	மட்டமான	50	20	10	80
மொத்தம்		100	60	40	200

இப்பொழுது பொது அறிவுக்கும் சமூக நடத்தைக்கும் இடையேயுள்ள கெழுவைக் காண்போம். இணைப்புக் கெழுவைக் காண, முதலில் χ^2-ன் மதிப்பைக் காணவேண்டும். கொடுக்கப்பட்டுள்ள இணைப்புப் பட்டியலின் கண்டறிந்த அலைவெண்களுக்கான எதிர்பார்க்கும் அலைவெண்கள் பின்வருமாறு அட்டவணைப்படுத்தப் பட்டுள்ளன.

		சமூக நடத்தை			மொத்தம்
		நல்ல	சுமாரான	கெட்ட	
பொது அறிவு	நல்ல	$\frac{90\times100}{200}=45$	$\frac{90\times60}{200}=27$	$\frac{90\times40}{200}=18$	90
	சுமாரான	$\frac{30\times100}{200}=15$	$\frac{30\times60}{200}=9$	$\frac{30\times40}{200}=6$	30
	மட்டமான	$\frac{80\times100}{200}=40$	$\frac{80\times60}{200}=24$	$\frac{80\times40}{200}=16$	80
மொத்தம்		100	60	40	200

இப்பொழுது

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

$$= \frac{(40-45)^2}{45} + \frac{(30-27)^2}{27} + \frac{(20-18)^2}{18} + \frac{(10-15)^2}{15}$$
$$+ \frac{(10-9)^2}{9} + \frac{(10-6)^2}{6} + \frac{(50-40)^2}{40} + \frac{(20-24)^2}{24}$$
$$+ \frac{(10-16)^2}{16}$$

$$= \frac{5^2}{45} + \frac{3^2}{27} + \frac{2^2}{18} + \frac{(-5)^2}{15} + \frac{1^2}{9} + \frac{4^2}{6} + \frac{10^2}{40}$$
$$+ \frac{(-4)^2}{24} + \frac{(-6)^2}{16}$$

$$= \frac{25}{45} + \frac{9}{27} + \frac{4}{18} + \frac{25}{15} + \frac{1}{9} + \frac{16}{6} + \frac{100}{40}$$
$$+ \frac{16}{24} + \frac{36}{16}$$

$$= 0{\cdot}56 + 0{\cdot}33 + 0{\cdot}22 + 1{\cdot}67 + 0{\cdot}11 + 2{\cdot}67 + 2{\cdot}50$$
$$+ 0{\cdot}67 + 2{\cdot}25$$

$= 10{\cdot}98$ ஆகும்.

எனவே

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

$$= \sqrt{\frac{10{\cdot}98}{10{\cdot}98 + 200}}$$

$$= \sqrt{\frac{10{\cdot}98}{210{\cdot}98}}$$

$$= \sqrt{0{\cdot}0520}$$

$= 0{\cdot}228$ ஆகும்.

பொது அறிவுக்கும் சமூக நடத்தைக்குமிடையேயுள்ள தொடர்பின் அளவான இணைப்புக் கெழு $C = 0 \cdot 228$ என அறிகிறோம்.

மாதிரி : மேலே கொடுக்கப்பட்ட மாதிரியில் $\chi^2 = 10 \cdot 98$ என்றும் $C = 0 \cdot 228$ என்றும் கண்டோம். பொது அறிவிற்கும் சமூக நடத்தைக்குமிடையே தொடர்பு இல்லை என்ற சூனிய எடுகோளையும் பொது அறிவிற்கும் சமூக நடத்தைக்குமிடையே தொடர்பு உள்ளது என்ற மாற்று எடுகோளையும் கொள்வோம். இப்பொழுது அட்டவணை C-யிலிருந்து $(k-1)(r-1) = (3-1)(3-1) = 4$ வரையற்ற பாகைகளுக்கு H_0-ன் கீழ் $\chi^2 \geqslant 10 \cdot 98$-க்கான நிகழ்தகவின் மதிப்பு $< 0 \cdot 05$ என அறிகிறோம். எனவே, H_0-ஐ $0 \cdot 05$ மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். அதாவது பொது அறிவிற்கும் சமூக நடத்தைக்குமிடையே தொடர்பு உள்ளது என்று உய்த்துணருகிறோம். அல்லது $C = 0 \cdot 228$ ஆனது மிகைத்தன்மை செய்யுமளவிற்கு பூச்சியத்திலிருந்து வேறுபட்டுள்ளது என முடிவு செய்கிறோம்.

ஸ்பியர்மேனின் தரக்கெழு r_s (Spearman Rank Correlation Coefficient r_s)

தரங்களைத் தழுவிய புள்ளியியல் அளவைகளில் ஸ்பியர்மேன் ஒட்டுறவுக் கெழுவும் ஒன்றாகும். இதனை r_s என்று குறிப்பிடுகிறோம். சில பண்புகளை அளக்க முடியாத நிலையில் அவற்றின் தர மதிப்புகள் பெறப்பட்டு, அத் தர மதிப்புகளுக்கிடையேயுள்ள தொடர்பு, அப் பண்புகளுக்கிடையேயுள்ள தொடர்பை எடுத்துக் காட்டுவதாக அமைகிறது. கண்டறிந்த விவரங்களை வரிசைப்படுத்தி, தரங்களைக் கொடுத்துத் தொடர்புப் போக்குக் கோடு நேர்கோடாக இருக்கும் என்ற அனுமானத்தின் அடிப்படையில் தொடர்புக்கு அளவையாக இந்த ஸ்பியர்மேன் தரத் தொடர்புக் கெழு அமைகிறது.

X, Y இரு மாறிகளுக்கேற்ப N நபர்களைத் தரப்படுத்தி $X_1, X_2, \ldots\ldots X_N ; Y_1, Y_2, \ldots\ldots Y_N$ என்று அமைப்போம். இவைகளுக்கிடையேயுள்ள தொடர்பு சிறந்ததாக இருந்தால் எல்லா i களுக்கும் $X_i = Y_i$ என்று இருக்கும். $X_i - Y_i = d_i$ எனக் கொண்டு, $d_i = 0$ எனில் இரு மாறிகளுக்கிடையேயுள்ள தொடர்பு சிறந்துள்ளது என்றும், d_i களின் மதிப்பு அதிகமாக இருந்தால் இரு மாறிகளுக்கிடையேயுள்ள தொடர்பு சிறந்ததாக இல்லை என்றும் ஊகமாக அறியலாம்.

இப்பொழுது ஸ்பியர்மேன் தரக்கெழுவைக் காண உதவும் சூத்திரத்தை வருவிப்போம். $x = X - \bar{X}$ என்றும் $y = Y - \bar{Y}$

என்றும் கொள்வோம். இதில் $\bar{X}$, $\bar{Y}$ என்பன X, Y ஆகிய மாறி களின் N மதிப்புகளுக்கான சராசரிகளாகும். சாதாரணமாகப் பயன்படுத்தப்படும் ஒட்டுறவுக் கெழு

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \, \Sigma y^2}} \text{ ஆகும்.} \tag{1}$$

X-களும், Y-களும் தரங்களாக இருந்தால் $r = r_s$ என்றிருக்கும். மேலும் $1, 2, \ldots\ldots, N$ என்ற முழு எண்களின் கூட்டுத் தொகையான $\Sigma X = \dfrac{N(N+1)}{2}$ என்றும் $1^2, 2^2 \ldots\ldots, N^2$ களின் கூட்டுத் தொகையான $\Sigma X^2 = \dfrac{N(N+1)(2N+1)}{6}$ என்றும் இருக்கும்.

$$\text{எனவே, } \Sigma x^2 = \Sigma (X - \bar{X})^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

$$= \frac{N(N+1)\ (2N+1)}{6} - \frac{N^2(N+1)^2}{4}$$

$$= \frac{N^3 - N}{12} \tag{2}$$

இதே போன்று $\Sigma y^2 = \dfrac{N^3 - N}{12}$ ஆகும்.

இப்பொழுது $d = x - y$

$$d^2 = (x - y)^2 = x^2 - 2xy + y^2$$

$$\Sigma d^2 = \Sigma x^2 + \Sigma y^2 - 2\,\Sigma xy$$

எனவே, கண்டறிந்தவை தரப்படுத்தப்பட்டால் சூத்திரம் (1)-ன்படி

$$r = \frac{\Sigma xy}{\sqrt{(\Sigma x)^2\,(\Sigma y)^2}} = r_s \text{ ஆகும்.}$$

$$\therefore \ \Sigma d^2 = \Sigma x^2 + \Sigma y^2 - 2r_s \sqrt{(\Sigma x)^2\ (\Sigma y)^2}$$

இவ்வாறாக,

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d^2}{2\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

இப்பொழுது,

$$\Sigma x^2 = \frac{N^3 - N}{12} = \Sigma y^2$$

எனப் பிரதியிட்டால்

$$r_s = \frac{\frac{N^3 - N}{12} + \frac{N^3 - N}{12} - \Sigma d^2}{2\sqrt{\left(\frac{N^3 - N}{12}\right)\left(\frac{N^3 - N}{12}\right)}}$$

$$= \frac{2\left(\frac{N^3 - N}{12}\right) - \Sigma d^2}{2\left(\frac{N^3 - N}{12}\right)}$$

$$= 1 - \frac{\Sigma d^2}{\frac{N^3 - N}{6}}$$

$$= 1 - \frac{6\,\Sigma d^2}{N^3 - N} \text{ ஆகும்.}$$

இவ்வாறாக $d = x - y = (X - \bar{X}) - (Y = \bar{Y})$

$$= (X - Y) - (\bar{X} - \bar{Y})$$

$$= X - Y \left\{ \begin{array}{l} \text{ஏனெனில் தரங்களுக்கு} \\ \bar{X} = \bar{Y} \text{ என்றிருப்பதால்} \end{array} \right\}$$

என்பதால் மேலுள்ள சூத்திரத்தை

$$r_s = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N^3 - N} \qquad (3)$$

என்று எழுதலாம். r_s-ன் மதிப்பு $-1 \leq r_s \leq +1$ என்று அமையும். சூத்திரம் (3)-ல் N என்பது கொடுக்கப்பட்ட இணை விவரங்களின் எண்ணிக்கையாகும். d_i என்பது தர மதிப்புகளின் வேறுபாடு ஆகும். கொடுக்கப்பட்டுள்ள இணை விவரங்களின் எண்ணிக்கையான N ஆனது 25-க்கும் குறைவாக இருந்தால்

இந்தச் சூத்திரத்தின் மூலம் இரு பண்புகளுக்கு இடையேயுள்ள தொடர்பினை அறியப் பயன்படுத்தலாம்.

மாதிரி : 12 மாணவர்கள் *X*, *Y* என்ற இரண்டு சோதனை களில் பெற்ற மதிப்பெண்கள் பின்வருமாறு :

மாணவர்கள்	X	Y
1	182	142
2	198	146
3	187	139
4	140	137
5	216	165
6	213	188
7	211	186
8	183	156
9	185	162
10	226	192
11	206	154
12	217	181

இப்பொழுது இரு சோதனைகளின் மதிப்புகளுக்கிடையேயுள்ள ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவைக் காண்போம்.

மாணவர்கள்	சோதனை தரமதிப்பு R_1	சோதனை தரமதிப்பு R_2	$d_i = R_1 - R_2$	d_i^2
1	2	3	−1	1
2	6	4	2	4
3	5	2	3	9
4	1	1	0	0
5	10	8	2	4
6	9	11	−2	4
7	8	10	−2	4
8	3	6	−3	9
9	4	7	−3	9
10	12	12	0	0
11	7	5	2	4
12	11	9	2	4
				$\Sigma d_i^2 = 52$

ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழு

$$r_s = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N^3 - N}$$

இங்கு $N = 12, \Sigma d_i^2 = 52$ ஆகும்.

எனவே $r_s = 1 - \frac{6(52)}{(12)^3 - 12}$

$$= 1 - \frac{312}{1728 - 12}$$

$$= 1 - \frac{312}{1716}$$

$$= 1 - \frac{26}{143}$$

$$= 1 - 0{\cdot}18$$

$= 0{\cdot}82$ ஆகும்.

மாதிரி 2 : நீச்சல் போட்டி, ஓட்டப் பந்தயம் ஆகியவற்றில் கலந்துகொண்ட பத்து நபர்களின் தரங்கள் கீழே கொடுக்கப் பட்டுள்ளன.

நபர்கள்	தரம்	
	ஓட்டப் பந்தயம்	நீச்சல் போட்டி
1	3	5
2	6	3
3	4	8
4	7	4
5	5	2
6	8	1
7	1	7
8	9	10
9	10	6
10	2	3

இவ்விவரங்களுக்கான ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவின் மதிப்பைக் காண்போம்.

நபர்கள்	தரம்		d_i	d_i^2
	ஓட்டப்பந்தயம்	நீச்சல்போட்டி		
1	3	5	−2	4
2	6	3	3	9
3	4	8	−4	16
4	7	4	3	9
5	5	2	3	9
6	8	1	7	49
7	1	7	−6	36
8	9	10	−1	1
9	10	6	4	16
10	2	3	−1	1
				$\Sigma d_i^2 = 150$

இங்கு $N = 10$, $\Sigma d_i^2 = 150$

எனவே $r_s = 1 - \frac{6\ \Sigma d_i^2}{N^3 - N}$

$$= 1 - \frac{6\ (150)}{(10)^3 - 10}$$

$$= 1 - \frac{900}{1000 - 10}$$

$$= 1 - \frac{900}{990}$$

$$= 1 - \frac{10}{11}$$

$$= 1 - 0{\cdot}91$$

$= 0{\cdot}09$ ஆகும்.

கவனிக்கவும் : கண்டறிந்த விவரங்களுக்குள் ஒத்தவை (ties) இருக்கும்போது, ஒத்த விவரங்களின் தர மதிப்புகளின் சராசரித் தர மதிப்பையே அவற்றின் தர மதிப்புகளாகக் கொள்ள வேண்டும். ஒத்தவைகளின் எண்ணிக்கை அதிகமாக இருப்பின் r_s-ன் மதிப்பைக் காண திருத்தக் காரணியத்தைக் (correction factor) கருதவேண்டும். X மாறியிலுள்ள ஒத்தவைகளுக்கான தரங்களின் விளைவானது Σx^2 ஐ $\frac{N^3 - N}{12}$ ஐ விடக் குறைவாக்குவதே யாகும். அதாவது ஒத்தவைகளுக்கான தரங்கள் X மாறியிலிருந்தால் $\Sigma x^2 < \frac{N^3 - N}{12}$. இந்த விலக்க வர்க்கங்களின்

கூட்டலுக்கான (sum of squares) திருத்தத்தைச் செய்ய, X மாறியில் ஏற்பட்டுள்ள ஒத்தவைகளின் எண்ணிக்கையைக் காணவேண்டும். அவ்வொத்தவைகளின் எண்ணிக்கையை t_x எனக் கொள்வோம். இப்பொழுது திருத்தக் காரணி $T_x = \frac{t_x^3 - t_x}{12}$ ஆகும்.

எனவே $\Sigma x^2 = \frac{N^3 - N}{12} - \Sigma T_x$. இதேபோன்று Y மாறிக்கான திருத்தக் காரணி $T_y = \frac{t_y^3 - t_y}{12}$ ஆகும்.

எனவே $\Sigma_y^2 = \frac{N^3 - N}{12} - \Sigma T_y$ ஆகும். இப்பொழுது திருத்தத்தைப் பெற்ற ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழு,

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d^2}{2\sqrt{(\Sigma x^2)(\Sigma y^2)}} \text{ ஆகும்.}$$

$$\text{இதில் } \Sigma x^2 = \frac{N^3 - N}{12} - T_x$$

$$\Sigma y^2 = \frac{N^3 - N}{12} - T_y \text{ ஆகும்.}$$

மாதிரி ; விரைவாக எழுதும் போட்டி, நினைவாற்றல் போட்டி ஆகியவற்றில் கலந்துகொண்ட 12 மாணவர்களின் மதிப்புகள் பின்வருமாறு :

மாணவர்கள்	மதிப்புகள்	
	விரைவாக எழுதும் போட்டி	நினைவாற்றல் போட்டி
A	10	52
B	10	56
C	11	49
D	11	47
E	13	75
F	14	98
G	15	96
H	16	66
I	17	72
J	18	102
K	18	64
L	12	91

விரைவாக எழுதும் தன்மைக்கும், நினைவாற்றலுக்கும் இடையே யுள்ள ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவின் மதிப்பைக் காண்போம்.

மாணவர்கள்	தரம்		d_i	d_i^2
	விரைவாக எழுதும் போட்டி	நினைவாற்றல் போட்டி		
A	1·5	3	−1·5	2·25
B	1·5	4	−2·5	6·25
C	3·5	2	1·5	2·25
D	3·5	1	2·5	6·25
E	5	8	−3·0	9·00
F	6	11	−5·0	25·00
G	7	10	−3·0	9·00
H	8	6	2·0	4·00
I	9	7	2·0	4·00
J	10·5	12	−1·5	2·25
K	10·5	5	−5·5	30·25
L	12	9	3·0	9·00
				$\Sigma d_i^2 = 109{\cdot}50$

இங்கு $N = 12$, $\Sigma d_i^2 = 109{\cdot}50$

எனவே,

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

இதில் $\Sigma x^2 = \frac{N^3 - N}{12} - \Sigma T_x$

$\Sigma y^2 = \frac{N^3 - N}{12} - \Sigma T_y$ ஆகும்.

இப்பொழுது

$$\Sigma x^2 = \frac{N^3 - N}{12} - \Sigma T_x$$

$$= \frac{12^3 - 12}{12} - \left\{ \left(\frac{2^3 - 2}{12}\right) + \left(\frac{2^3 - 2}{12}\right) + \left(\frac{2^3 - 2}{12}\right) \right\}$$

$$= \frac{12\,(12^2-1)}{12} - \left\{\frac{6}{12}+\frac{6}{12}+\frac{6}{12}\right\}$$

$$= (144-1) - \left\{\frac{1}{2}+\frac{1}{2}+\frac{1}{2}\right\}$$

$$= 143 - 1{\cdot}5$$

$$= 141{\cdot}5$$

நினைவாற்றல் போட்டிகளின் மதிப்புகளில் ஒத்தவை இல்லை. எனவே $\Sigma T_y = 0$ ஆகும்.

அதனால் $$\Sigma y^2 = \frac{N^3-N}{12} - \Sigma T_y$$

$$= \frac{12^3-12}{12} - 0$$

$$= \frac{12\,(12^2-1)}{12}$$

$$= 144 - 1$$

$$= 143$$

இவ்வாருக ஒத்தவைகளுக்குத் திருத்தப்பட்ட $\Sigma x^2 = 141{\cdot}5$, $\Sigma y^2 = 143$ ஆகும்,

எனவே, $$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

$$= \frac{141{\cdot}5 + 143 - 109{\cdot}5}{2\sqrt{(141{\cdot}5)(143)}}$$

$$= \frac{284{\cdot}5 - 109{\cdot}5}{2\sqrt{(141{\cdot}5)(143)}}$$

$$= \frac{175}{2\sqrt{(141{\cdot}5)(143)}}$$

$$= \frac{175}{2\sqrt{20234{\cdot}5}}$$

$$= \frac{175}{2 \times (142{\cdot}24)}$$

$$= \frac{175}{284{\cdot}48}$$

$= 0{\cdot}616$ ஆகும்.

ஸபியர்மேன் தர ஒட்டுறவுக் கெழுவிற்கு மிகைத்தன்மைச் சோதனை (Testing the Significance of Spearman's Rank Correlation Coefficient)

r_s-ன் மதிப்பைக் காணப் பயன்படுத்தப்பட்ட மதிப்புகள் ராண்டம் முறையில் ஏதோ ஒரு முழுமைத் தொகுதியிலிருந்து தேர்ந்தெடுக்கப்பட்டவை எனில் அம் மதிப்புகளைப் பயன்படுத்தி இரு மாறிகள் அம் முழுமைத் தொகுதியில் தொடர்பு உடையன வாக உள்ளனவா என்பதை அறியலாம். அதாவது ஆய்வில் உள்ள இரு மாறிகள் முழுமைத் தொகுதியில் தொடர்பற்றவை என்றும் கண்டறிந்த r_s-ன் மதிப்பு பூச்சியத்திலிருந்து தற்செயலாக வேறு படுகிறது என்ற சூனிய எடுகோளையும் சோதனை செய்ய விரும்புகிறோம்.

இப்பொழுது சிறு கூறுகளுக்கான சோதனையைப் பற்றிக் காண்போம்.

சூனிய எடுகோள் உண்மையானது என்று கொள்வோம். அதாவது X, Y ஆகிய மாறிகளுக்கிடையே முழுமைத் தொகுதியில் தொடர்பு இல்லை என்று கொள்வோம், X, Y ஆகியவைகளின் மாதிரி மதிப்புகள் அம் முழுமைத் தொகுதியிலிருந்து ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டவை எனில் கொடுக்கப்பட்ட Y மதிப்புகளின் தர வரிசைக்கு X மதிப்புகளின் எந்தத் தர வரிசையும் அந்த X மதிப்புகளின் மற்றத் தர வரிசைக்குச் சமவாய்ப்புடைய தாக இருக்கும்.

மேலும், கொடுக்கப்பட்ட X மதிப்புகளின் தர வரிசைக்கு, Y மதிப்புகளின் இயலும் எல்லாவிதமான வரிசையும் சம வாய்ப் புடையனவாக இருக்கும். N நபர்களைக் கொண்ட X மதிப்புகள் $N!$ இயலுமான தரங்களைப் பெற்றுக் கொடுக்கப்பட்ட Y மதிப்புகளின் எந்தத் தரங்களுக்கும் தொடர்புடையதாக இருக்கலாம். இத் தொடர்புகள் சம வாய்ப்புடையனவாக இருப்பதால் X மதிப்பு களின் தரமும் கொடுக்கப்பட்ட Y மதிப்புகளின் தரத்திற்கு ஏற்ப நிகழ்வதற்கான நிகழ்தகவு $\frac{1}{N!}$ ஆகும். ஒவ்வொரு தர வரிசையில் இயலுமான Y மதிப்புகளின் எவ்விதமான தரங்களுக்கும் தொடர் புடைய r_s-ன் மதிப்பு இருக்கும். H_0-ன் கீழ் எந்தக் குறிப்பிட்ட r_s-ன் மதிப்பும் நிகழ்வதற்கான நிகழ்தகவு அம்மதிப்பு ஏற்படுவதற் கான பல்வேறு வரிசை மாற்றத்திற்கு (permulation) விகிதச் சமமாக (proportional) அமைகிறது.

$$r_s = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N^3 - N}$$

என்ற சூத்திரத்தைப் பயன்படுத்தி $N = 2$-க்கு r_s-ன் முடியுமான இரண்டு மதிப்புகள் $+1$, -1 என்பதை அறிகிறோம். ஒவ்வொன்றும் H_0-ன் கீழ் $\frac{1}{2}$ என்ற நிகழ்தகவு மதிப்பைப் பெறுகின்றன. இதே போன்று $N = 3$-க்கு r_s-ன் முடியுமான மதிப்புகள் $-1, -\frac{1}{2}, +\frac{1}{2}, 1$ ஆகும். எனவே H_0-ன் கீழ் அம் மதிப்புகளுக்கு ஏற்ற முறையான நிகழ்தகவு மதிப்புகள் $\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}$ ஆகும். அட்டவணை P-ல் r_s-ன் தீர்வு கட்டமான மதிப்புகள் இம்முறையில் கண்டுபிடிக்கப்பட்டு அட்டவணைப்படுத்தப்பட்டுள்ளன. அதில் $N = 4$-லிலிருந்து 30 வரை $p = 0{\cdot}05$, $p = 0{\cdot}1$-க்கேற்ற H_0-ன் கீழ் தொடர்புடைய நிகழ்தகவு மதிப்புகள் கொடுக்கப்பட்டுள்ளன. இது ஒருமுனைச் சோதனைக்கான நிகழ்தகவு மதிப்புகளைப் பெற்ற அட்டவணையாகும். கண்டறிந்த r_s-ன் மதிப்பு அட்டவணை P-யிலுள்ள மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் கண்டறிந்த மதிப்பு அம் மிகைத்தன்வம மட்டத்தில் மிகைத் தன்மை செய்கிறது என்று உய்த்துணருகிறோம். அதாவது H_0-ஐ மறுக்கிறோம்.

மாதிரி : மாதிரி 1-ல் $N = 12$ என்றும் $r_s = 0{\cdot}82$ என்றும் கண்டோம். அட்டவணை P-யிலிருந்து $N = 12$-க்கும் $\alpha = 0{\cdot}01$-க்கும் ஏற்ற மதிப்பு $0{\cdot}712$ ஆகும். கண்டறிந்த r_s-ன் மதிப்பான $0{\cdot}82$ அட்டவணை மதிப்பைவிட அதிகமாக உள்ளதால் கண்டறிந்த r_s-ன் மதிப்பு $\alpha = 0{\cdot}01$ என்ற மட்டத்தில் மிகைத்தன்மை செய்கிறது என்று உய்த்துணருகிறோம். அதாவது ஆய்விலுள்ள இரு மாறிகளுக்கிடையே முழுமைத்தொகுதியில் தொடர்பு இல்லை என்ற சூனிய எடுகோளான H_0-ஐ $\alpha = 0{\cdot}01$ என்ற மிகைத் தன்மை மட்டத்தில் மறுக்கிறோம்.

மாதிரி : மாதிரி 3-ல் $N = 12$, $r_s = 0{\cdot}6$ எனக் கண்டோம். இப்பொழுது $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தில் அட்டவணை P-யிலுள்ள மதிப்பு $0{\cdot}712$ ஆகும். கண்டறிந்த r_s-ன் மதிப்பு அட்டவணையின் மதிப்பைவிடக் குறைவாக உள்ளதால் H_0-ஐ $\alpha = 0{\cdot}01$ என்ற மிகைத்தன்மை மட்டத்தில் ஏற்றுக்கொள்கிறோம். ஆனால் $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் அட்டவணை P-யிலுள்ள மதிப்பு $0{\cdot}506$ ஆகும். எனவே, கண்டறிந்த r_s-ன் மதிப்பு அட்டவணை மதிப்பைவிட அதிகமாக உள்ளதால் H_0-ஐ $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

இப்பொழுது பெரிய கூறுகளுக்கு ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவுக்கான மிகைத்தன்மைச் சோதனையைக் காண்போம்.

$N \geqslant 10$ எனில் அக் கூறுகளைப் பெரிய கூறுகள் என்கிறோம். கண்டறிந்த r_s-ன் மிகைத்தன்மையை H_0-ன் கீழ் சோதனை செய்ய

$$t = r_s \sqrt{\frac{N-2}{1-r_s^2}}$$

என்ற புள்ளியியல் அளவையால் சோதனை செய்கிறோம். N-ன் அளவு பெரியதாக இருந்தால் மேலுள்ள புள்ளியியல் அளவை t ஆனது $(N-2)$ வரையற்ற பாகையைப் பெற்ற ஸ்டூடன்ட்-t பரவலாக அமைகிறது. அட்டவணை B-ஐப் பயன்படுத்தி $(N-2)$ வரையற்ற பாகைக்கும், α என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற கண்டறிந்த t-ன் மதிப்பைவிட அதிகமாக இருப்பதற்கான தீர்வுகட்டமான t-ன் மதிப்புகளைப் பெறலாம். கண்டறிந்த t-ன் மதிப்பு அட்டவணை B-ல். $N-2$ வரையற்ற பாகைக்கும், α என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் H_0-ஐ α என்ற மிகைத்தன்மை மட்டத்தில் ஒருமுனைச் சோதனைக்கோ அல்லது இருமுனைச் சோதனைக்கோ மறுக்கிறோம்.

மாதிரி : திறமைத் தேர்வில் 15 விற்பனையாளர்கள் பெற்ற மதிப்பெண்கள், அவ் விற்பனையாளர்களின் முதலாண்டு விற்பனை மதிப்புகள் ஆகியவற்றைப் பற்றிய விவரங்கள் கீழ்க்கண்ட அட்டவணையில் கொடுக்கப்பட்டுள்ளன.

விற்பனை யாளர்கள்	திறமைத்தேர்வு மதிப்பெண்கள்	விற்பனை மதிப்புகள்
1	36	32
2	15	21
3	22	33
4	20	20
5	34	31
6	23	26
7	27	25
8	33	30
9	25	22
10	31	29
11	38	34
12	26	36
13	28	28
14	18	27
15	17	24

இவ் விவரங்களைக் கொண்டு முழுமைத் தொகுதியில் திறமைத் தேர்விற்கும் விற்பனைக்கும் இடையே தொடர்புள்ளதா என்பதை ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவுக்கான மிகைத்தன்மைச் சோதனை செய்து அறிவோம். ஆய்வில் கொள்ளப்பட்ட திறமைத் தேர்வு, விற்பனை ஆகிய இரு மாறிகளுக்குமிடையே முழுமைத் தொகுதியில் தொடர்பு இல்லை என்ற சூனிய எடுகோளைக் கருதுவோம். இப்பொழுது இரு மாறிகளுக்குமான தரங்களையும் அவைகளுக்கிடையேயுள்ள முறையான வேறுபாட்டையும் அட்டவணைப் படுத்துவோம்.

விற்பனையாளர்கள்	தரம்		d_i	d_i^2
	X திறமைத் தேர்வு	Y விற்பனை		
1	2	4	−2	4
2	15	14	1	1
3	11	3	8	64
4	12	15	−3	9
5	3	5	−2	4
6	10	10	0	0
7	7	11	−4	16
8	4	6	−2	4
9	9	13	−4	16
10	5	7	−2	4
11	1	2	−1	1
12	8	1	7	49
13	6	8	−2	4
14	13	9	4	16
15	14	12	2	4
				$\Sigma d_i^2 = 196$

இப்பொழுது

$$r_s = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N^3 - N}$$

இங்கு $N = 15$, $\Sigma d_i^2 = 196$ ஆகும்.

எனவே

$$
\begin{aligned}
r_s &= 1 - \frac{6(196)}{(15)^3 - 15} \\
&= 1 - \frac{1176}{15(15^2 - 1)} \\
&= 1 - \frac{1176}{15 \times 224} \\
&= 1 - \frac{1176}{3360} \\
&= 1 - 0{\cdot}35 \\
&= 0{\cdot}65 \text{ ஆகும்.}
\end{aligned}
$$

இங்கு $N = 15$ என்பது பெரிய கூறு என்பதால் பெரிய கூறுகளுக்கான சோதனைப் புள்ளியியல் அளவை

$$
\begin{aligned}
t &= r_s \sqrt{\frac{N-2}{1-r_s^2}} \\
&= 0{\cdot}65 \sqrt{\frac{15-2}{1-(0{\cdot}35)^2}} \\
&= 0{\cdot}65 \sqrt{\frac{13}{1-0{\cdot}1225}} \\
&= 0{\cdot}65 \sqrt{\frac{13}{0{\cdot}8775}} \\
&= 0{\cdot}65 \sqrt{14{\cdot}81} \\
&= 0{\cdot}65 \times 3{\cdot}84 \\
&= 2{\cdot}496 \text{ ஆகும்.}
\end{aligned}
$$

இது $N-2 = 15-2 = 13$ வரையற்ற பாகையைப் பெற்ற ஸ்டூடன்ட்-t பரவலாக அமைகிறது. இப்பொழுது அட்டவணை B-யிலிருந்து ஒருமுனைச் சோதனைக்கு $\alpha = 0{\cdot}05$ என்ற மிகைத் தன்மை மட்டத்திற்கும், 13 என்ற வரையற்ற பாகைக்கும் ஏற்ற தீர்வுகட்டமான மதிப்பு 1·771 ஆகும். கண்டறிந்த t-ன் மதிப்பான 2·496 ஆனது அட்டவணையின் மதிப்பைவிட அதிகமாக உள்ளதால் H_0-ஐ $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம். அதாவது திறமைத்தேர்வு, விற்பனை ஆகிய இரு மாறிகளுக்குமிடையே முழுமைத் தொகுதியில் தொடர்பு இல்லை என்ற சூனிய எடுகோளை $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

குறிப்பு : அட்டவணை P-ல், ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவான r_s-ன் தீர்வுகட்டமான மதிப்புகள் $N = 30$ வரை கொடுக்கப்பட்டுள்ளன. Journal of the American Statistical Association (September 1972, volume 67, number 339) என்ற பத்திரிகையில் $N = 4\ (1)\ 50\ (2)\ (100)$ வரை $\alpha\ (1)$ என்ற ஒரு முனைச் சோதனைக்கான r_s-ன் தீர்வுகட்டமான மதிப்புகளும், $\alpha\ (2)$ என்ற இருமுனைச் சோதனைக்கான r_s-ன் தீர்வுகட்டமான மதிப்புகளும், பக்கம் 578-ல், $\alpha = 0{\cdot}50, 0{\cdot}20, 0{\cdot}10, 0{\cdot}05, 0{\cdot}02$ $0{\cdot}01, 0{\cdot}005, 0{\cdot}002, 0{\cdot}001$ என்ற மிகைத்தன்மை மட்டங்களுக்கு ஜெரால்டு ஹெச். ஸார் என்பவரால் கொடுக்கப் பட்டுள்ளன.

கெண்டால் தர ஒட்டுறவுக் கெழு τ (Kendall Rank Correlation Coefficient)

ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுவைக் காணப் பயன் படுத்தப்பட்ட அதே விவரங்களுக்கு கெண்டால் தர ஒட்டுறவுக் கெழுவைக் காணலாம். இரு கூறுகளின் தரங்களுக்கிடையே யுள்ள தொடர்பு அல்லது ஒட்டுறவை அளவிடும் இந்தச் சோதனை யின் அளவையை τ எனக் குறிப்பிடுகிறோம். r_s-ஐப் போலவே τ-ஐயும் பயன்படுத்தி மிகைத்தன்மைச் சோதனை செய்யலாம். தர ஒட்டுறவுக் கெழுவான τ-ஐப் பகுதி ஒட்டுறவுக் கெழுவாகப் (partial correlation coefficient) பொதுப்படையாக அமைக்கலாம். ஆனால் r_s-ஐ இவ்வாறு பகுதி ஒட்டுறவுக் கெழுவாகப் பொதுப் படையாக அமைக்க முடியாது.

இச் சோதனையின் அம்சத்தை ஓர் உதாரணம் மூலம் காண் போம். அ, ஆ, இ, ஈ என்ற நான்கு கட்டுரைகளை X, Y என்ற இரு நீதிபதிகள் தரப்படுத்தி அமைத்த விவரங்கள் பின்வருமாறு :

கட்டுரை	அ	ஆ	இ	ஈ
நீதிபதி X	3	4	2	1
நீதிபதி Y	3	1	4	2

நீதிபதி X-ன் தரங்களை 1, 2, 3, 4 என்று இயற்கையான வரிசை யில் அமைத்து அதற்கு ஏற்ற நீதிபதி Y-ன் தரங்களைப் பின் வருமாறு அமைப்போம்.

கட்டுரை	ஈ	இ	அ	ஆ
நீதிபதி X	1	2	3	4
நீதிபதி Y	2	4	3	1

இப்பொழுது X, Y ஆகிய நீதிபதிகளின் நீதிகளுக்கிடையேயுள்ள தொடர்பின் அளவைக் காண இருக்கிறோம். நீதிபதி X-ன் தரங்கள் இயற்கையான வரிசையில் அமைந்துள்ளன என்பதால் நீதிபதி Y-ன் தரங்களில் எத்தனை ஜோடிகள் இயற்கை வரிசையிலும், இயற்கை வரிசைக்குப் புறம்பாகவும் அமைந்துள்ளன என்பதை அறியலாம். மேலுள்ள அட்டவணையில் (2, 4) என்ற ஜோடியானது சரியான இயற்கை வரிசையில் அமைந்துள்ளது. ஒரு ஜோடி சரியான இயற்கை வரிசையில் அமைந்திருந்தால் + 1 என்றும், இயற்கை வரிசைக்குப் புறம்பாக அமைந்திருந்தால் —1 என்றும் அமைத்துக்கொள்வோம். இவ்வாறாக முதல் ஜோடியான (2,4) ஆனது இயற்கையான வரிசையில் அமைந்துள்ளது. எனவே, இந்த ஜோடிக்கு + 1 என்று அளிக்கிறோம். அடுத்தபடியாக (2,3) என்ற ஜோடியும் இயற்கை வரிசையில் அமைந்துள்ளது. எனவே, இந்த ஜோடிக்கு +1 என்று அளிக்கிறோம். இறுதியாக (2, 1) என்ற ஜோடி இயற்கை வரிசைக்குப் புறம்பாக அமைந்துள்ளது. எனவே, இந்த ஜோடிக்கு —1 என்று அளிக்கிறோம். எனவே, தரம் 2-ஐ உள்ளடக்கிய எல்லா ஜோடிகளுக்குமான மதிப்புகள்

$$(+1) + (+1) + (-1) = +1 \text{ ஆகும்.}$$

அடுத்தபடியாகத் தரம் 4-ஐ உள்ளடக்கிய எல்லா ஜோடிகள் (4, 3), (4, 1) ஆகும். இதில் (4, 3)-ம் (4, 1)-ம் இயற்கை வரிசைக்குப் புறம்பாக அமைந்துள்ளன. எனவே, தரம் 4-ஐ உள்ளடக்கிய எல்லா ஜோடிகளுக்குமான மதிப்புகள்

$$(-1) + (-1) = -2 \text{ ஆகும்.}$$

இறுதியாகத் தரம் 3-ஐ உள்ளடக்கிய ஒரே ஒரு ஜோடி (3, 1) ஆகும். இது இயற்கை வரிசைக்குப் புறம்பாக அமைந்துள்ளது. எனவே, இந்த ஜோடிக்கு —1 என்று அளிக்கிறோம்.

எல்லா மதிப்புகளின் மொத்தம் : $(+1) + (-2) + (-1) = -2$ ஆகும். முடியுமான உச்ச அளவிலான மொத்தம் $= \binom{4}{2} = 6$ ஆகும். τ என்ற கெண்டால் தர ஒட்டுறவுக்

கெழுவை எல்லா மதிப்புகளின் மொத்தத்திற்கும், முடியுமான உச்ச அளவிலான மொத்தத்திற்கும் உள்ள விகிதமாகக் குறிக்கிறேம். அதாவது

$$\tau = \frac{\text{கணக்கிடப்பட்ட மொத்தம்}}{\text{முடியுமான உச்ச அளவிலான மொத்தம்}}$$

$$= \frac{-2}{6}$$

$$= -0\cdot33 \text{ ஆகும்.}$$

$\tau = -0\cdot33$ என்பது X, Y ஆகிய இரு நீதிபதிகள் அளித்த தரங்களுக்கிடையேயுள்ள தொடர்பின் அளவாகும். பொதுவாக

$$\tau = \frac{\text{கணக்கிடப்பட்ட மதிப்பு}}{\text{இயலுமான உச்ச மதிப்பு}}$$

என்று வரையறை செய்கிறேம். இதில் இயலுமான உச்ச மதிப்பு N_{C_2} எனக் கொள்கிறேம். உண்மையான மதிப்பு $+1$, -1 ஆகியவைகளின் மொத்தத் தொகை என்பதால் இதனை S என்று குறிப்பிட்டு

$$\tau = \frac{S}{N_{C_2}}$$

$$= \frac{S}{\frac{N(N-1)}{2}}$$

என்று அமைக்கிறேம். இதில் N என்பது X, Y ஆகியவைகளில் தரப்படுத்தப்பட்ட நபர்களின் அல்லது பொருள்களின் எண்ணிக்கையாகும்.

S-ன் மதிப்பை எளிய முறையில் பின்வருமாறு காணலாம். மேலுள்ள உதாரணத்தில் நீதிபதி X-ன் இயற்கை வரிசையிலான தரங்களுக்கு ஏற்ற நீதிபதி Y-ன் தரங்கள் 2,4,3,1 ஆகும். இதில் முதலில் அமைந்துள்ள தரமான 2-க்கு வலப் புறத்தில் இதைவிட அதிகமாக உள்ள தரங்களின் எண்ணிக்கை இரண்டு என்றும் (அதாவது 4, 3 ஆகியவை) குறைவாக உள்ள தரங்களின் எண்ணிக்கை ஒன்று என்றும் (அதாவது 1) அறிகிறேம். எனவே, தரம் 2-க்கு ஏற்ற S-க்கான மதிப்பு $(+2\ -1) = +1$ ஆகும். இதே போன்று அடுத்த தரமான 4-க்கு வலப் புறத்தில் இதைவிட அதிகமான தரங்கள் எதுவும் இல்லை. எனவே, அவைகளின் எண்ணிக்கை 0 ஆகும். ஆனல், தரம் 4-க்கு வலப் புறத்தில்

இதைவிடக் குறைவான தரங்களின் எண்ணிக்கை இரண்டாகும் (3, 1 ஆகியவை). எனவே, தரம் 4-க்கு ஏற்ற S-க்கான மதிப்பு $(0-2) = -2$ ஆகும். அடுத்தபடியாகத் தரம் 3-க்கு வலப் புறத்தில் இதைவிட அதிகமான தரங்கள் எதுவுமில்லை. ஆனால், அதைவிடக் குறைவான தரங்களின் எண்ணிக்கை ஒன்று ஆகும். எனவே, தரம் 3-க்கு ஏற்ற S-க்கான மதிப்பு $(0-1) = -1$ ஆகும்.

எனவே, உண்மை மதிப்பு $S = (+1) + (-2) + (-1)$

$= -2$ ஆகும்.

$$\text{எனவே } \tau = \frac{S}{\frac{1}{2}N(N-1)}$$

$$= \frac{-2}{\frac{1}{2}(4)(4-1)}$$

$$= \frac{-2}{2 \times 3}$$

$$= -\frac{1}{3}$$

$= -0{\cdot}33$ ஆகும்.

மாதிரி : 10 மாணவர்கள் A, B என்ற இரு தேர்வுகளில் பெற்ற மதிப்பெண்கள் பின்வருமாறு :

மாணவர்கள்	1	2	3	4	5	6	7	8	9	10
தேர்வு A	32	28	46	54	59	62	44	39	42	68
தேர்வு B	48	72	64	38	46	51	49	65	59	67

மேலுள்ள விவரங்களுக்கு கெண்டால் தர ஒட்டுறவுக் கெழுவைக் காண்போம். தேர்வு A, தேர்வு B ஆகியவைகளுக்கான தரங்கள் பின்வருமாறு :

மாணவர்கள்	a	b	c	d	e	f	g	h	i	j
தேர்வு A தரம்	9	10	5	4	3	2	6	8	7	1
தேர்வு B தரம்	8	1	4	10	9	6	7	3	5	2

τ-ன் மதிப்பைக் காண, தேர்வு A-ன் தரங்களை இயற்கை வரிசையில் அமைத்து அவைகளுக்கேற்ற தேர்வு B-ன் தரங்களைக் கீழ்க்கண்டவாறு அமைப்போம்.

மாணவர்கள்	j	f	e	d	c	g	i	h	a	b
தேர்வு A தரம்	1	2	3	4	5	6	7	8	9	10
தேர்வு B தரம்	2	6	9	10	4	7	5	3	8	1

தேர்வு B-ல் முதலில் அமைந்துள்ள தரமான 2-க்கு வலப் புறத்தில் இதைவிட அதிகமாக உள்ள தரங்களின் எண்ணிக்கை 8 ஆகும்; குறைவாக உள்ள தரங்களின் எண்ணிக்கை 1 ஆகும்; இதேபோன்று அடுத்த தரமான 6-க்கு வலப் புறத்தில் இதைவிட அதிகமாக உள்ள தரங்களின் எண்ணிக்கை 4 ஆகும்; குறைவாக உள்ள தரங்களின் எண்ணிக்கையும் 4 ஆகும். இவ்வாறே அடுத்த தரமான 9-க்கு வலப் புறத்தில் இதைவிட அதிகமாக உள்ள தரங்களின் எண்ணிக்கை 1 ஆகும்; குறைவாக உள்ள தரங்களின் எண்ணிக்கை 6 ஆகும். இம் முறையில் கணக்கிட்டால்

$$S = (8-1)+(4-4)+(1-6)+(0-6)+(3-2)$$
$$+(1-3)+(1-2)+(1-1)+(0-1)$$
$$= 7+0-5-6+1-2-1+0-1$$
$$= -7 \text{ ஆகும்.}$$

எனவே $$\tau = \frac{S}{\frac{1}{2}N(N-1)}$$
$$= \frac{-7}{\frac{1}{2}(10)(10-1)}$$
$$= \frac{-7}{\frac{1}{2}\times 10\times 9}$$
$$= \frac{-7}{45}$$
$$= -0{\cdot}155$$

$\tau = -0{\cdot}155$ என்பது தேர்வு A, தேர்வு B ஆகியவைகளுக் கிடையேயுள்ள தொடர்பின் அளவைத் தெரிவிக்கிறது.

கவனிக்கவும் : X-லோ அல்லது Y-லோ இரண்டு அல்லது அதற்கு மேற்பட்ட கண்டறிந்தவை ஒத்தவைகளாக (tied observations) இருந்தால், ஒத்தவைகளுக்கு முன்பு கண்டவாறு சராசரியிலான தரங்களை அளிக்கிறோம். இம்மாதிரியான ஒத்தவைகளால் ஏற்படும் விளைவைத் திருத்தி

$$\tau = \frac{S}{\sqrt{\{\frac{1}{2}N(N-1)-T_X\}}\ \sqrt{\{\frac{1}{2}N(N-1)-T_Y\}}}$$

என்று அமைத்துக்கொள்ளுகிறோம். இதில் $T_X = \frac{1}{2}\Sigma t_x\ (t_x-1)$ ஆகும்; t_x என்பது X மாறியிலுள்ள ஒத்தவைகளின் எண்ணிக்கை யாகும்; $T_Y = \frac{1}{2}\ \Sigma t_y\ (t_y - 1)$ ஆகும்; t_y என்பது Y மாறியிலுள்ள ஒத்தவைகளின் எண்ணிக்கையாகும்.

மாதிரி : A, B என்னும் இரு உளவியல் சோதனையில் பங்கு கொண்ட 15 மாணவர்கள் பெற்ற மதிப்பெண்களின் விவரம் பின்வருமாறு :

மாணவர்கள்	சோதனை A	சோதனை B
l	180	110
m	203	98
n	188	118
o	193	104
p	176	112
q	174	124
r	158	119
s	197	95
t	177	94
u	138	97
v	126	110
w	160	94
x	151	126
y	186	119
z	187	117

இவ்விவரங்களைக் கொண்டு சோதனை A, சோதனை B ஆகியவைகளுக்கிடையேயுள்ள கெண்டால் தர ஒட்டுறவுக் கெழுவைக் காண்போம்.

இப்பொழுது சோதனை A, சோதனை B ஆகியவைகளுக்கான தரங்களைக் கீழ்க்காணும் அட்டவணையில் அமைப்போம்.

மாணவர்கள்		l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
தரங்கள்	சோதனை A	7	1	4	3	8	10	12	2	9	14	15	11	13	5	6
	சோதனை B	8·5	11	5	10	7	2	3·5	13	14·5	12	8·5	14·5	1	3·5	6

இப்பொழுது சோதனை A-ன் தரங்களை இயற்கை வரிசையில் அமைத்து அவைகளுக்கேற்ற சோதனை B-ன் தரங்களைக் கீழ்க்கண்டவாறு அமைப்போம்.

மாணவர்கள்		m	s	o	n	y	z	l	p	t	q	w	r	x	u	v
தரங்கள்	சோதனை A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	சோதனை B	11	13	10	5	3·5	6	8·5	7	14·5	2	14·5	3·5	1	12	8·5

வழக்கமான முறையில் கண்டால்

$$S = (4-10)+(2-11)+(3-9)+(7-4)+(7-2)+(6-3) +(3-4)+(4-3)+(0-5)+(4-1)+(0-4)+(2-1) +(2-0)+(0-1)$$

$$= (-6)+(-9)+(-6)+3+5+3+(-1)+1+(-5) +3+(-4)+1+2+(-1)$$

$$= 18-32$$

$$= -14$$

X : சோதனை A-ன் மதிப்புகளில் ஒத்தவை எதுவுமில்லை எனவே $T_X = 0$. Y : சோதனை B-ன் மதிப்புகளில் இரண்டு $3 \cdot 5$-ம் இரண்டு $14 \cdot 5$-ம் கொண்ட தரங்கள் உள்ளன. எனவே

$$T_Y = \frac{1}{2} \Sigma t_y (t_y - 1)$$
$$= \frac{1}{2} \{2(2-1) + 2(2-1) + 2(2-1)\}$$
$$= \frac{1}{2} (2+2+2)$$
$$= 3 \text{ ஆகும்.}$$

இவ்வாறாக $T_X = 0$, $T_Y = 3$, $S = -14$, $N = 15$ ஆகியவைகளைக் கொண்டு,

$$\tau = \frac{S}{\sqrt{\{\frac{1}{2} N(N-1) - T_X\}} \sqrt{\{\frac{1}{2} N(N-1) - T_Y\}}}$$

$$= \frac{-14}{\sqrt{\frac{1}{2}(15)(15-1) - 0} \sqrt{\frac{1}{2}(15)(15-1) - 3}}$$

$$= \frac{-14}{\sqrt{105-0} \sqrt{105-3}}$$

$$= \frac{-14}{\sqrt{105 \times 102}}$$

$$= \frac{-14}{103 \cdot 5}$$

$= -0 \cdot 13$ எனக் காண்கிறோம்.

τ-ஐயும், r_s-ஐயும் ஒப்பிடுதல் (Comparison of τ and r_s)

கொடுக்கப்பட்ட அதே விவரங்களுக்குக் கணக்கிடப்படும் τ, r_s ஆகியவைகளின் மதிப்புகள் சமமாக இருக்கமாட்டா. τ-ஐயும்

r_s-ஐயும் காணப் பயன்படுத்தப்படும் அடிப்படையான அளவு கோல்கள் வெவ்வேறு வகைப்பட்டவை. மேலும் அவைகளை எண்ணளவியலாக (numerically) ஒன்றுக்கொன்று நேராக ஒத்துப் பார்க்கமுடியாது. அதாவது X, Y ஆகிய மாறிகளுக்கிடையேயுள்ள r_s என்ற ஒட்டுறவையும், X, Z ஆகிய மாறிகளுக்கிடையேயுள்ள τ என்ற ஒட்டுறவு அளவையையும் கண்டால் r_s, τ ஆகியவைகளை ஒத்துப் பார்க்கக்கூடிய அளவைகளில்லை என்பதால் X ஆனது Y-க்கோ அல்லது Z-க்கோ மிக நெருங்கிய தொடர்புடையதாக இருக்கும் என்று கூறமுடியாது.

எனினும் இரண்டு கெழுக்களும் ஒரே விவரங்களைப் பயன் படுத்தி முழுமைத் தொகுதியில் அக் கூறுகளுக்கிடையே தொடர்பு இருக்கிறதா இல்லையா என்பதை அறிவிக்கும் திறன் படைத்ததாய் உள்ளது. அதாவது τ, r_s ஆகியவைகளின் மாதிரிப் பரவல்கள், கொடுக்கப்பட்ட விவரங்களைக் கொண்டு ஒரு குறிப்பிட்ட மிகைத் தன்மை மட்டத்தில் சூனிய எடுகோளை மறுக்கும் அல்லது ஏற்றுக் கொள்ளும்.

τ-க்கான மிகைத்தன்மைச் சோதனை (Testing the Significance of τ)

X, Y ஆகிய மாறிகள் தொடர்பற்று இருக்கும் முழுமைத் தொகுதியிலிருந்து ஒரு ராண்டம் மாதிரியைத் தேர்ந்தெடுப் போம். X, Y ஆகியவைகளின் பேரில் இந்த மாதிரியின் உறுப்பு கள் தர வரிசைப்படுத்தப்படுகின்றன. பிறகு கொடுக்கப்பட்ட X தர வரிசைகளுக்கு, Y-ன் முடியுமான எல்லாத் தர வரிசைகள் சம வாய்ப்பைப் பெற்றிருக்கும். அதாவது கொடுக்கப்பட்ட X தரங்களின் வரிசைகளுக்கு Y-ன் முடியுமான எந்த ஒரு தர வரிசை யும், மற்ற முடியுமான தர வரிசைகளைப்போல் நிகழ்வது சமவாய்ப் புடையதாக இருக்கும். உதாரணமாக X தரங்களை $1, 2, 3 \ldots \ldots \ldots, N$ என்று இயற்கை வரிசையில் அமைப்போம். இந்த X தரங்களின் வரிசைக்கு $N!$ முடியுமான Y தரங்களின் வரிசைகள் H_0-ன் கீழ் சமவாய்ப்புடையதாக இருக்கும். எனவே எந்த ஒரு குறிப்பிட்ட Y-ன் தரங்களின் வரிசையும் H_0-ன் கீழ் $\frac{1}{N!}$ என்ற நிகழ்தகவு மதிப்பைப் பெற்றதாயிருக்கும். Y-ன் முடியுமான $N!$ தரங்களில், ஒவ்வொன்றுக்கும் தொடர்புடைய τ-ன் மதிப்பு இருக்கும். τ-ன் முடியுமான மதிப்புகள் -1 லிருந்து $+1$ வரை அமைந்திருக்கும். இவைகளை அலைவெண் பரவலாக (frequency distribution) அமைக்கலாம். உதாரணமாகப் பின்வரும் அலைவெண் பரவலைக் கருதுவோம்.

τ	அலைவெண்
$-1 \cdot 0$	1
$-0 \cdot 7$	4
$-0 \cdot 4$	6
0	8
$0 \cdot 4$	6
$0 \cdot 7$	4
$1 \cdot 0$	1

இவ்வலைவெண் பரவலில் உள்ள ஒவ்வொரு τ-ன் மதிப்பும், ஒவ்வொரு குறிப்பிட்ட அலைவெண்களை H_0-ன் கீழ் பெற்றுள்ளது. இப்பொழுது இவை $N = 5$-க்கு H_0-ன் கீழ் நிகழ்வதற்கான நிகழ்தகவு பின்வருமாறு:

τ	H_0-ன் கீழ் அமைந்துள்ள அலைவெண்	H_0-ன் கீழ் நிகழ்வதற்கான நிகழ்தகவு
$-1 \cdot 0$	1	$\frac{1}{120}$
$-0 \cdot 7$	4	$\frac{4}{120}$
$-0 \cdot 4$	6	$\frac{6}{120}$
0	8	$\frac{8}{120}$
$0 \cdot 4$	6	$\frac{6}{120}$
$0 \cdot 7$	4	$\frac{4}{120}$
$1 \cdot 0$	1	$\frac{1}{120}$

இதேபோன்று N-ன் மற்ற மதிப்புகளுக்கு நிகழ்தகவு அட்டவணைகளை அமைக்கமுடியும். ஆனால் N-ன் மதிப்பு அதிகமாக ஆக இம்முறை சற்றுக் கடினமானதாகிறது.

$N \geqslant 8$ எனில் τ-ன் மாதிரிப் பரவலானது இயல்நிலைப் பரவலாக அமைந்துள்ளது என்று கெண்டால் என்பவரால் கண்டறியப்

பட்டுள்ளது. எனவே N-ன் பெரிய மதிப்புகளுக்கு, H_0-ன் கீழ் τ-ன் மதிப்பு நிகழ்வதற்கான இயல்நிலை நிகழ்தகவு மதிப்புகள் அட்டவணை A-ல் கொடுக்கப்பட்டுள்ளன. எனினும் $N \leqslant 10$ எனில் அட்டவணை Q-ல் H_0-ன் கீழ் கண்டறிந்த S-ன் மதிப்பைவிட அதிக மாக இருப்பதற்கான ஒருமுனைச் சோதனைக்கேற்ற நிகழ்தகவு மதிப்புகள் கொடுக்கப்பட்டுள்ளன. இவ்வாறு சிறிய கூறுகளுக்கு இரு கூறுகளின் தரங்களுக்கிடையேயுள்ள கண்டறிந்த தொடர்புக் கான α என்ற மிகைத்தன்மை மட்டத்தை S-ன் மதிப்புக்கேற்ற அட்டவணை Q-வுள்ள மதிப்பைக் கண்டு அறியலாம். $p \leqslant \alpha$ எனில் H_0-ஐ மறுக்கிறோம். உதாரணமாக $N = 9$, $S = 12$ என்க. அட்டவணை Q-விலிருந்து H_0-ன் கீழ் $N = 9$-க்கு $S \geqslant 11$ என்றிருப் பதற்கான நிகழ்தகவு $p = 0 \cdot 130$ ஆகும்.

N-ன் அளவு 10-ஐ விட அதிகமாக இருந்தால் τ ஆனது சராசரி $= \mu = 0$ஐயும்

$$\text{திட்டவிலக்கம்} = \sigma = \sqrt{\frac{2(2N+5)}{9N(N-1)}}$$

என்று பெற்ற இயல்நிலைப் பரவலாக அமைகிறது. அதாவது

$$z = \frac{\tau - \mu}{\sigma}$$

$$= \frac{\tau}{\sqrt{\dfrac{2(2N+5)}{9N(N-1)}}}$$

சராசரி $= 0$, மாறுபாடு $= 1$ ஆகியவைகளைப் பெற்ற தோராய மான இயல்நிலைப் பரவலாக அமைகிறது. அட்டவணை A-லிருந்து z-ஐ மிகைத்தன்மை செய்யும் மட்டங்களை அறிந்து H_0-ஐ சோதனை செய்யலாம்.

மாதிரி : இரு நீதிபதிகள் 10 மாணவர்களுக்கு ஒரு போட்டி யில் அளித்த மதிப்பெண்கள் பின்வருமாறு :

X	4	6	12	17	9	14	10	8	21	5	24	23
Y	13	16	15	18	11	5	6	22	19	20	25	23

இரு மாறிகளும் ஏதோ ஒரு முழுமைத் தொகுதியிலிருந்து எடுக்கப் பட்ட ராண்டம் மாதிரி எனில், இந்த இரு மாறிகளும் அந்த முழுமைத் தொகுதியில் தொடர்புடையனவாக உள்ளனவா என் பதை τ-க்கான மிகைத்தன்மைச் சோதனை மூலம் அறிவோம்.

இங்கு $N = 12$ ஆகும். முதலில் கெண்டால் தர ஒட்டுறவுக் கெழுவான τ-ஐக் கணக்கிடுவோம்.

தரங்கள்	X	12	10	6	4	8	5	7	9	3	11	1	2
	Y	9	7	8	6	10	12	11	3	5	4	1	2

இப்பொழுது X-ன் இயற்கையான தர வரிசைக்கு ஏற்ற Y-ன் தரங்கள் பின்வருமாறு :

தரங்கள்	X	1	2	3	4	5	6	7	8	9	10	11	12
	Y	1	2	5	6	12	8	11	10	3	7	4	9

$$S = (11-0) + (10-0) + (7-2) + (6-2) + (0-7) + (3-3) + (0-5) + (0-4) + (3-0) + (1-1) + (1-0)$$

$$= 11 + 10 + 5 + 4 - 7 + 0 - 5 - 4 + 3 + 0 + 1$$

$= 22$ ஆகும்.

எனவே $\tau = \dfrac{S}{\frac{1}{2}N(N-1)}$

$$= \frac{22}{\frac{1}{2}(12)(12-1)}$$

$$= \frac{22}{\frac{1}{2} \times 12 \times 11}$$

$$= \frac{22}{66}$$

$= 0 \cdot 33$ ஆகும்.

இங்கு $N > 10$ என்பதால் τ-க்கான மிகைத்தன்மைச் சோதனையை மேற்கொள்ள

$$z = \frac{\tau}{\sqrt{\dfrac{2(2N+5)}{9N(N-1)}}}$$

என்ற புள்ளியியல் அளவையைப் பயன்படுத்துகிறோம். இது சராசரி $= 0$, மாறுபாடு $= 1$ ஆகியவைகளைப் பெற்ற இயல்நிலைப் பரவலாக அமைகிறது.

இப்பொழுது,

$$z = \frac{\tau}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}}$$

$$= \frac{0 \cdot 33}{\sqrt{\frac{2\ (24+5)}{9(12)(12-1)}}}$$

$$= \frac{0 \cdot 33}{\sqrt{\frac{2 \times 29}{9 \times 12 \times 11}}}$$

$$= \frac{0 \cdot 33}{\sqrt{\frac{58}{1188}}}$$

$$= \frac{0 \cdot 33}{\sqrt{0 \cdot 049}}$$

$$= \frac{0 \cdot 33}{0 \cdot 22}$$

$= 1 \cdot 5$ ஆகும்.

அட்டவணை A-யிலிருந்து H_0-ன் கீழ் $z \geqslant 1 \cdot 5$-க்கான நிகழ்தகவு $p = 0 \cdot 0668$ ஆகும். எனவே H_0-ஐ $\alpha = 0 \cdot 0668$ என்ற மிகைத் தன்மை மட்டத்தில் மறுத்து, இரு மாறிகளும் அந்த முழுமைத் தொகுதியில் தொடர்புடையனவாக உள்ளன என்று உய்த்துணருகிறோம்.

கெண்டாலின் பகுதி தர ஒட்டுறவுக் கெழு $\tau_{xy \cdot z}$ (Kendall's Partial Rank Correlation Coefficient)

தர ஒட்டுறவுக் கெழுவைப் பயன்படுத்தி இரண்டு மாறிகளுக் கிடையேயுள்ள தொடர்பின் அளவைக் கண்டோம். இப்பொழுது இரண்டுக்கு மேற்பட்ட மாறிகளுக்கிடையேயுள்ள தொடர்பைப் பற்றிக் காண்போம்.

X, *Y* ஆகிய மாறிகளின் உறவை ஆயும்பொழுது, அவ்விரு மாறிகளுக்கிடையேயான ஒட்டுறவுக் கெழுவின் மதிப்பு $0 \cdot 7$

என்போம். இம் மதிப்பைக் கொண்டு Y-ன் மதிப்பு அதிகரிக்க, அதிகரிக்க X-ன் மதிப்பு அதிகரிக்கும் என்று உட்பொருள் விளக்கம் பெறுகிறோம். X-ன் மதிப்பில் மாற்றம் ஏற்படக்கூடிய காரணிகள் யாவும் நிலையாக இருக்கும்போது X-க்கும் Y-க்கும் இடையேயுள்ள ஒட்டுறவுக்கெழு $+ 0{\cdot}70$ எனில் மேற்கூறிய பொருள் விளக்கம் பொருத்தமானதாகிறது. X-ன் மதிப்புகளைப் பாதிக்கும் காரணிகள் நிலையாக இல்லாமல் மாறுபடும்போது மேற்கூறிய பொருள் விளக்கம் உண்மை நிலைக்குப் புறம்பாக அமையக்கூடும். ஏனெனில் மாறி Y ஆனது, X மாறியுடன் தொடர்பு அற்று இருந்த போதிலும் இவ்விரு மாறிகள் ஒவ்வொன்றும் வேறொரு மாறியான Z உடன் உறவுகொண்டிருந்தால் X-க்கும் Y-க்கும் இடையேயுள்ள ஒட்டுறவுக் கெழு கணிசமான மதிப்பைப் பெற்று அமையலாம். ஆகவே, இரு மாறிகளின் ஒட்டுறவுக் கெழு மதிப்பை மட்டும் கொண்டு அவைகளின் உறவு பற்றிய முழுமையான கருத்தைப் பெற முற்படுவது தவறானதாகும்.

இரு மாறிகளின் உறவைப்பற்றி அறிய முற்படும்போது அம் மாறிகளுக்குப் பொதுவான காரணிகள் யாவையும், மாறுபடாமல் நிலையாக அமையும்படி செய்தல் இன்றியமையாததாகும். அவ்வாறு பொதுவான காரணிகள் யாவும் நிலையாக அமையப் பெறாவிட்டால் பொதுக்காரணிகளின் விளைவை ஒவ்வொரு மாறியிலிருந்தும் நீக்கிய பின்பு கிடைக்கும் மதிப்புகளைக் கொண்டு பெறப்படும் ஒட்டுறவுக் கெழுவானது, மாறிகளின் உண்மைத் தொடர்பை நன்கு வெளிப்படுத்தும்படி அமையும். இவ்வாறு பெறப்படும் ஒட்டுறவுக் கெழுவைப் பகுதி ஒட்டுறவுக் கெழு (Partial Correlation Coefficient) என அழைக்கிறோம்.

சுட்டுறுப்பைச் சாராத முறைகளுக்குப் பகுதி ஒட்டுறவைக் காண கெண்டாலின் பகுதி தர ஒட்டுறவுக் கெழுவைப் பயன்படுத்துகிறோம். கொடுக்கப்படும் விவரமானது வரிசை அளவுகோலுக்கு அளவிடும்படியாக இருக்குமானால் சுட்டுறுப்பைச் சாராத பகுதி தர ஒட்டுறவுக் கெழுவைக் காணமுடியும். இங்கு முழுமைத் தொகுதியின் வடிவத்தைப் பற்றி எந்தவிதமான தற்கோளையும் (assumptions) நாம் மேற்கொள்வதில்லை.

உதாரணமாக l, m, n, o என்ற 4 நபர்களைக் கொண்ட X, Y, Z என்ற மாறிகளின் தரங்களைக் கருதுவோம். Z ஆனது நிலையாக இருக்கும்போது X, Y ஆகிய மாறிகளுக்கிடையேயுள்ள ஒட்டுறவைப் பற்றிக் காண்போம். இந்த மாறிகளுக்கான தரங்களை உதாரணத்திற்காகப் பின்வருமாறு அமைப்போம்.

நபர்கள்	*l*	*m*	*n*	*o*
Z-ன் தரங்கள்	1	2	3	4
X-ன் தரங்கள்	3	1	2	4
Y-ன் தரங்கள்	2	1	3	4

இப்பொழுது எந்த மாறியின் முடியுமான எல்லா ஜோடித் தரங்களையும் (4_{C_2}) வழிகளில் பெறலாம். *Z*-ன் தரங்களை இயற்கையான வரிசையில் அமைத்த பிறகு முடியுமான ஒவ்வொரு *X*-ன் ஜோடித் தரங்களையும், *Y*-ன் தரங்களையும், *Z*-ன் தரங்களையும் கவனிப்போம். *X*, *Y*, *Z* ஆகியவைகளின் ஒவ்வொரு ஜோடியிலும் உள்ள தரங்களில் முதலிலுள்ள தரம், இரண்டாம் தரத்தை விடக் குறைவாக இருந்தால் + குறியையும், முதலிலுள்ள தரம் இரண்டாம் தரத்தைவிட அதிகமாக இருந்தால் — குறியையும் அமைப்போம். உதாரணமாக (*l*, *m*) என்ற ஜோடியில் *X*-க்கான ஜோடித் தரம் (3, 1) ஆகும். இதில் 3 ஆனது 1-ஐ விட அதிகமாக உள்ளதால் — குறியை அளிக்கிறோம்; *Y*-க்கான ஜோடித் தரம் (2, 1) ஆகும். இதில் 2 ஆனது 1-ஐ விட அதிகமாக உள்ளதால் — குறியை அளிக்கிறோம். இவ்வாறு *l*, *m*, *n*, *o* ஆகியவைகளில் முடியுமான $(4_{C_2}) = 6$ ஜோடிகளுக்கான *X*, *Y*, *Z* ஆகிய மாறிகளுக்கான தரங்களைக் கீழ்க்கண்டவாறு அமைப்போம்.

ஜோடி	(*l*, *m*)	(*l*, *n*)	(*l*, *o*)	(*m*, *n*)	(*m*, *o*)	(*n*, *o*)
Z	+	+	+	+	+	+
X	—	—	+	+	+	+
Y	—	+	+	+	+	+

மேலுள்ள அட்டவணையிலுள்ள விவரங்களைக் கொண்டு *Z*-ன் குறிகளோடு ஒத்த *X*-க்கான குறிகளின் அலைவெண்களையும், *Z*-ன் குறிகளோடு ஒத்த *Y*-க்கான குறிகளின் அலைவெண்களையும், *Z*-ன் குறிகளோடு ஒத்திராத *X*-க்கான குறிகளின் அலைவெண்

களையும், Z-ன் குறிகளோடு ஒத்திராத Y-க்கான குறிகளின் அலைவெண்களையும் (2 × 2) இணைப்புப் பட்டியலில் கீழ்க் கண்டவாறு அமைப்போம்.

	Z-ன் குறிகளோடு ஒத்த Y-ன் குறிகள்	Z-ன் குறிகளோடு ஒத்திராத Y-ன் குறிகள்	மொத்தம்
Z-ன் குறிகளோடு ஒத்த X-ன் குறிகள்	4	0	4
Z-ன் குறிகளோடு ஒத்திராத X-ன் குறிகள்	1	1	2
மொத்தம்	5	1	6

பொதுவாக மேலுள்ள அட்டவணையைக் கீழ்க்கண்டவாறு அமைப்போம்.

	Z-ன் குறிகளோடு ஒத்த Y-ன் குறிகள்	Z-ன் குறிகளோடு ஒத்திராத Y-ன் குறிகள்	மொத்தம்
Z-ன் குறிகளோடு ஒத்த X-ன் குறிகள்	s	t	$s+t$
Z-ன் குறிகளோடு ஒத்திராத X-ன் குறிகள்	u	v	$u+v$
மொத்தம்	$s+u$	$t+v$	$\binom{N}{2}$

Z ஆனது நிலையாக இருக்கும்பொழுது X, Y ஆகிய மாறிகளுக்கிடையேயுள்ள கெண்டாலின் பகுதி தர ஒட்டுறவுக் கெழுவை

$$\tau_{xy \cdot z} = \frac{sv - ut}{\sqrt{(s+t)(u+v)(s+u)(t+v)}}$$

என்று வரையறுக்கிறோம். மேற்கூறிய உதாரணத்தில் $N=4$ ஆகும். எனவே

$$\tau_{xy \cdot z} = \frac{4(1)-(0)(1)}{\sqrt{(4)(2)(5)(1)}}$$

$$= \frac{4}{\sqrt{40}}$$

$$= \frac{4}{6 \cdot 32}$$

$= 0 \cdot 63$ ஆகும்.

மேலுள்ள $\tau_{xy \cdot z}$-ன் சூத்திரத்தை பைக் கெழு (Phi-coefficient) என்றும் கூறுகிறோம். சோதனை உருப்படிகளானவை 'ஆம் அல்லது இல்லை', 'வெற்றி அல்லது தோல்வி' என்றிருக்கும் விவரங்களுக்கு அடிக்கடி பயன்படுத்தப்படுகின்றன.

மேலும் $\tau_{xy \cdot z}$ ஐ $\tau_{xy \cdot z} = \sqrt{\frac{\chi^2}{N}}$ என்றும் எழுதலாம்.

இதில் $\chi^2 = \frac{N\,(sv-ut)}{(s+t)\,(u+v)\,(s+u)\,(t+v)}$ ஆகும்.

N-ன் அளவு பெரியதாக இருந்தால் மேற்சொன்ன முறையின் மூலம் $\tau_{xy \cdot z}$-ன் மதிப்பைப் பெறுவது சற்றுக் கடினமானதாகிறது. கெண்டால் என்பவர்

$$\tau_{xy \cdot z} = \frac{\tau_{xy} - \tau_{zy} \cdot \tau_{xz}}{\sqrt{(1-\tau_{zy}^2)\,(1-\tau_{xz}^2)}}$$

என்று கண்டுள்ளார். இதில் $\tau_{xy}, \tau_{zy}, \tau_{xz}$ ஆகியவை X, Y; Z, Y; X, Z ஆகிய மாறிகளுக்கிடையேயுள்ள கெண்டாலின் ஒட்டுறவுக் கெழுக்களாகும்.

மாதிரி : X, Y, Z ஆகிய மாறிகளின் 12 மதிப்புகள் கீழே கொடுக்கப்பட்டுள்ளன.

X	Y	Z
86	46	4
102	50	4
91	43	5
44	41	5
120	69	7
117	92	8
115	90	9
87	60	10
89	66	11
130	96	12
110	58	12
121	85	16

இம் மாறிகளுக்கிடையேயுள்ள $\tau_{xy \cdot z}$ என்ற கெண்டால் பகுதி தர ஒட்டுறவுக் கெழுவைக் காண்போம்.

இங்கு $N = 12$ என்பதால்,

$$\tau_{xy \cdot z} = \frac{\tau_{xy} - \tau_{zy} \cdot \tau_{xz}}{\sqrt{(1 - \tau_{zy}^2)(1 - \tau_{xz}^2)}}$$

என்ற சூத்திரத்தைப் பயன்படுத்துவோம். இப்பொழுது τ_{xy}, τ_{zy}, τ_{xz} என்ற கெண்டாலின் தர ஒட்டுறவுக் கெழுக்களைக் காண்போம். கீழுள்ள அட்டவணையில் X, Y, Z ஆகிய மாறிகளின் தரங்கள் அமைக்கப்பட்டுள்ளன.

தரங்கள்		
X	**Y**	**Z**
3	2	1·5
4	6	1·5
2	5	3·5
1	1	3·5
8	10	5·0
11	9	6·0
10	8	7·0
6	3	8·0
7	4	9·0
12	12	10·5
5	7	10·5
9	11	12·0

முதலில் τ_{xy} ஐக் காண்போம். இதற்கு X, Y ஆகியவைகளின் தரங்களைக் கொண்டு S_{xy}-ன் மதிப்பைக் கண்ட பிறகு

$$\tau_{xy} = \frac{S_{xy}}{\frac{1}{2} N (N-1)}$$

என்ற சூத்திரத்தைக் கொண்டு காணவேண்டும்.

தரங்கள்	
X	Y
3	2
4	6
2	5
1	1
8	10
11	9
10	8
6	3
7	4
12	12
5	7
9	11

இப்பொழுது இயற்கை வரிசையிலமைந்த X-ன் தரங்களுக்கேற்ற Y-ன் தரங்களைக் கீழ்க்கண்டவாறு அமைத்து S_{xy}-ன் மதிப்பைக் காண்போம்.

தரங்கள்	X	1	2	3	4	5	6	7	8	9	10	11	12
	Y	1	5	2	6	7	3	4	10	11	8	9	12

$$S_{xy} = (11-0)+(7-3)+(9-0)+(6-2)+(5-2)+(6-0)+(5-0)+(2-2)+(1-2)+(2-0)+(1-0)$$

$$= 11+4+9+4+3+6+5+0-1+2+1$$

$$= 44$$

எனவே $r_{xy} = \dfrac{S_{xy}}{\frac{1}{2}N(N-1)}$

$$= \frac{44}{\frac{1}{2}\times 12\times(12-1)}$$

$$= \frac{44}{6\times 11}$$

$$= \frac{44}{66}$$

$$= \frac{2}{3}$$

$= 0 \cdot 67$ ஆகும்.

அடுத்தபடியாக Z, Y ஆகியவைகளின் தரங்களைக்கொண்டு S_{zy}-ன் மதிப்பைக் கண்ட பிறகு τ_{zy}-ன் மதிப்பைக் காணவேண்டும்.

தரங்கள்	
Y	Z
2	1·5
6	1·5
5	3·5
1	3·5
10	5·0
9	6·0
8	7·0
3	8·0
4	9·0
12	10·5
7	10·5
11	12·0

இப்பொழுது இயற்கை வரிசையில் அமைந்த Y-ன் தரங்களுக்கேற்ற Z-ன் தரங்களைக் கீழ்க்கண்டவாறு அமைத்தபிறகு S_{zy}-ன் மதிப்பைக் காண்போம்.

தரங்கள்													
	Y	1	2	3	4	5	6	7	8	9	10	11	12
	Z	3·5	1·5	8·0	9·0	3·5	1·5	10·5	7·0	6·0	5·0	12·0	10·5

இப்பொழுது,

$$S_{zy} = (8-2)+(9-0)+(4-5)+(3-5)+(6-1)+(6-0)$$
$$+(1-3)+(2-2)+(2-1)+(2-0)+(0-1)$$
$$= 6+9-1-2+5+6-2+0+1+2-1$$
$$= +29-6$$
$$= 23$$

இங்கு Z-ன் தரங்களில் ஒத்தவை உள்ளன என்பதால்

$$\tau_{zy} = \frac{S_{zy}}{\sqrt{\{\frac{1}{2}N(N-1)-T_z\}}\sqrt{\{\frac{1}{2}N(N-1)-T_y\}}}$$

என்ற சூத்திரத்தைப் பயன்படுத்துகிறோம். Z-ன் தரங்களில் 1·5 ஆனது இரு முறையும், 3·5 ஆனது இரு முறையும், 10·5 ஆனது இரு முறையும் ஏற்பட்டுள்ளன. எனவே

$$\begin{aligned} T_z &= \tfrac{1}{2}\Sigma t_z(t_z-1) \\ &= \tfrac{1}{2}\{2(2-1)+2(2-1)+2(2-1)\} \\ &= \tfrac{1}{2}\{2+2+2\} \\ &= \tfrac{6}{2} \\ &= 3 \text{ ஆகும்.} \end{aligned}$$

Y-ன் தரங்களில் ஒத்தவை இல்லை என்பதால் $T_y = 0$ என்று அளிக்கிறோம்.

எனவே,

$$\begin{aligned} \tau_{zy} &= \frac{S_{zy}}{\sqrt{\{\frac{1}{2}N(N-1)-T_z\}}\sqrt{\{\frac{1}{2}N(N-1)-T_Y\}}} \\ &= \frac{23}{\sqrt{\frac{1}{2}(12)(12-1)-3}\sqrt{\frac{1}{2}(12)(12-1)-0}} \\ &= \frac{23}{\sqrt{\frac{1}{2}\times 12\times 11-3}\sqrt{\frac{1}{2}\times 12\times 11-0}} \\ &= \frac{23}{\sqrt{66-3}\sqrt{66-0}} \\ &= \frac{23}{\sqrt{63}\sqrt{66}} \\ &= \frac{23}{\sqrt{4158}} \\ &= 0{\cdot}36 \text{ ஆகும்.} \end{aligned}$$

இறுதியாக X, Z ஆகிய மாறிகளின் தரங்களைக் கொண்டு S_{xy}-ன் மதிப்பைக் கண்ட பிறகு τ_{xz}-ன் மதிப்பைக் காண்போம்.

தரங்கள்	
X	Z
3	1·5
4	1·5
2	3·5
1	3·5
8	5·0
11	6·0
10	7·0
6	8·0
7	9·0
12	10·5
5	10·5
9	12·0

இப்பொழுது இயற்கை வரிசையிலமைந்த X-ன் தரங்களுக்கு ஏற்ற Z-ன் தரங்களைக் கீழ்க்கண்டவாறு அமைத்த பிறகு S_{xz}-ன் மதிப்பைக் காண்போம்.

தரங்கள்													
	X	1	2	3	4	5	6	7	8	9	10	11	12
	Z	3·5	3·5	1·5	1·5	10·5	8·0	9·0	5·0	12·0	7·0	6·0	10·5

இப்பொழுது

$$S_{xz} = (8-2)+(8-2)+(8-0)+(8-0)+(1-5)+(3-3)$$
$$+(2-3)+(4-0)+(0-3)+(1-1)+(1-0)$$
$$= 6+6+8+8-4+0-1+4-3+0+1$$
$$= +33-8$$
$$= 25$$

இங்கு Z-ன் தரங்களில் ஒத்தவை (ties) உள்ளன என்பதால்

$$\tau_{xz} = \frac{S_{xz}}{\sqrt{\{\frac{1}{2}N(N-1)-T_z\}}\ \sqrt{\{\frac{1}{2}N(N-1)-T_X\}}}$$

என்ற சூத்திரத்தைப் பயன்படுத்துகிறோம்.

Z-ன் தரங்களில் தரம் 1·5 ஆனது இரு முறையும், 3·5 ஆனது இரு முறையும், 10·5 ஆனது இரு முறையும் ஏற்பட்டுள்ளன.

எனவே

$$\begin{aligned} T_z &= \tfrac{1}{2}\, \Sigma t_z\,(t_z - 1) \\ &= \tfrac{1}{2}\,\{2(2-1) + 2(2-1) + 2(2-1)\} \\ &= \tfrac{1}{2}\,(2 + 2 + 2) \\ &= 3 \end{aligned}$$

X-ன் தரங்களில் ஒத்தவை ஏதுமில்லை என்பதால் $T_X = 0$ என்று அறிகிறோம். எனவே,

$$\begin{aligned} \tau_{xz} &= \frac{S_{xz}}{\sqrt{\{\tfrac{1}{2} N(N-1) - T_z\}}\ \sqrt{\{\tfrac{1}{2} N(N-1) - T_X\}}} \\ &= \frac{25}{\sqrt{\{\tfrac{1}{2}(12)(12-1) - 3\}}\ \sqrt{\{\tfrac{1}{2}(12)(12-1) - 0}} \\ &= \frac{25}{\sqrt{66-3}\ \sqrt{66-0}} \\ &= \frac{25}{\sqrt{63}\ \sqrt{66}} \\ &= \frac{25}{\sqrt{4158}} \\ &= 0{\cdot}39 \text{ ஆகும்.} \end{aligned}$$

இவ்வாறாக,

$$\begin{aligned} \tau_{xy\cdot z} &= \frac{\tau_{xy} - \tau_{zy}\,.\,\tau_{xz}}{\sqrt{(1-\tau_{zy}^2)\,(1-\tau_{xz}^2)}} \\ &= \frac{0{\cdot}67 - (0{\cdot}36)\,(0{\cdot}39)}{\sqrt{\{1-(0{\cdot}36)^2\}\,\{1-(0{\cdot}39)^2\}}} \\ &= \frac{0{\cdot}67 - 0{\cdot}1404}{\sqrt{(1-0{\cdot}1296)\,(1-0{\cdot}1521)}} \\ &= \frac{0{\cdot}5296}{\sqrt{(1-0{\cdot}1296)\,(1-0{\cdot}1521)}} \end{aligned}$$

$$= \frac{0 \cdot 5296}{\sqrt{(0 \cdot 8704)\ (0 \cdot 8479)}}$$

$$= \frac{0 \cdot 5296}{\sqrt{0 \cdot 73801216}}$$

$$= \frac{0 \cdot 5296}{0 \cdot 8590}$$

$= 0 \cdot 62$ ஆகும்.

மிகைத்தன்மைச் சோதனை (Test of Significance)

$r_{xy \cdot z}$-ன் மதிப்புக்கு விளக்கம் காண $r_{xy \cdot z}$-ஐ χ^2-க்குத் தொடர்பு படுத்தலாம். அதாவது

$$r_{xy \cdot z}^2 = \frac{\chi^2}{N}$$

அல்லது $\chi^2 = N \cdot r_{xy \cdot z}^2$ ஆகும்.

எனவே, கணக்கிடப்பட்ட χ^2-ஐ, அட்டவணை C-யிலுள்ள α என்ற மிகைத்தன்மை மட்டத்திற்கு ஏற்ற தீர்வுகட்டமான மதிப்புடன் ஒப்பிட்டுச் சோதனை செய்து அறியலாம். மேலே கண்ட மாதிரிக் கணக்கில் $r_{xyz} = 0 \cdot 62$ எனக் கண்டோம். எனவே

$$\chi^2 = N \cdot r_{xy \cdot z}^2$$
$$= 12 \times (0 \cdot 62)^2$$
$$= 12 \times 0 \cdot 3844$$
$$= 4 \cdot 6128 \text{ ஆகும்.}$$

இந்த χ^2 ஆனது $(2-1)(2-1) = 1$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாகும். $\alpha = 0 \cdot 05$ என்ற மிகைத் தன்மை மட்டத்தில் 1 வரையற்ற பாகைக்கேற்ற அட்டவணை C-யிலுள்ள மதிப்பு $3 \cdot 84$ ஆகும்.

எனவே, கண்டறிந்த χ^2-ன் மதிப்பானது $\alpha = 0 \cdot 05$ என்ற மிகைத்தன்மை மட்டத்திற்கும் 1 வரையற்ற பாகைக்கும் ஏற்ற அட்டவணை C-யிலுள்ள மதிப்பான $3 \cdot 84$-ஐ விட அதிகமாக உள்ளதால் Z நிலையாக இருக்கும்பொழுது X-க்கும் Y-க்கும் தொடர்பு இல்லை என்ற சூனிய எடுகோளை மறுத்து, அவைகளுக்கிடையே குறிப்பிடத்தக்க தொடர்பு உள்ளது என்று உய்த்துணருகிறோம்.

கெண்டாலின் இணக்கத் தொடர்புக் கெழு W (Kendall's Coefficient of Concordance)

இதற்கு முந்தின பகுதிகளில் N உருப்படிகளின் (individuals) தரங்களைப் பெற்ற இரு மாறிகளுக்கு இடையேயுள்ள ஒட்டுறவின் அளவைகளைக் கண்டோம். இப்பொழுது N உருப்படிகளின் தரங்களைப் பெற்ற பல மாறிகளுக்கிடையேயுள்ள தொடர்பின் அளவையைப் பற்றிக் காண்போம்.

வெவ்வேறு தரங்களைப் பெற்ற k மாறிகளுக்கிடையேயுள்ள தொடர்பைக் கெண்டாலின் இணக்கக் கெழுவான W-ஐப் பயன் படுத்தி அறியலாம். மொத்தத்தில் k மாறிகளுக்கிடையேயுள்ள தொடர்பை அறிய அந்த k மாறிகளில் ஒவ்வொரு முடியுமான ஜோடிகளுக்கு இடையேயுள்ள ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுக்களான r-களைக் கண்ட பிறகு அக் கெழுக்களின் சராசரியைக் கணக்கிடுகிறோம். இம்முறையைக் கைப்பற்றுவதால் $\binom{k}{2}$ தரக் கெழுக்களைக் காணவேண்டியுள்ளது. k-ன் அளவு சிறியதாக இருந்தாலன்றி இம் முறையைக் கையாள்வது மிகவும் கடின மாகிறது. கெண்டாலின் இணக்கத் தொடர்புக் கெழுவான W ஆனது k மாறிகளின் முடியுமான எல்லா ஜோடிகளின் ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுக்களான r_s களுடன் ஒரு படித்தான தொடர்பைப் பெற்று அமைகிறது. $\binom{k}{2}$ வழிகளில் பெறப்படும் ஒவ்வொரு ஜோடிகளுக்கிடையே உள்ள ஸ்பியர்மேன் தர ஒட்டுறவுக் கெழுக்களுக்குக் காணப்படும் சராசரியை $r_{s\,\text{சராசரி}}$ எனக் குறிப்பிட்டு, கெண்டால் என்பவர்,

$$r_{s\,\text{சராசரி}} = \frac{kW-1}{k-1} \text{ என்று நிறுவியுள்ளார்.}$$

இதில் W என்பது கெண்டாலின் இணக்கத் தொடர்புக் கெழுவாகும். N நபர்களின் தரங்களைக் கொண்ட k மாறி களுடைய நிரல்களின் கூட்டுத் தொகையை R_j என்று குறிப்பிடு கிறோம். பிறகு R_j-க்களின் கூட்டுத்தொகையை N-ஆல் வகுத்தால் R_j-க்களின் சராசரி கிடைக்கிறது. பிறகு ஒவ்வொரு R_j-க்களையும், இப்பொழுது கண்ட R_j-க்களின் சராசரிக்கு விலக்கம் (deviation) காண்கிறோம். இதனை s என்று குறிப்பிட்டு இவ்விலக்கத்தின் வர்க்கத்தைக் (square) காண்கிறோம்.

$$\text{அதாவது } s = \sum \left(R_j - \frac{\Sigma R_j}{N}\right)^2 \text{ ஆகும்.}$$

$$\text{இப்பொழுது, } W = \frac{s}{\frac{1}{12} k^2 (N^3 - N)} \text{ என்று } W\text{-க் காண்கிறோம்.}$$

உதாரணமாக *A*, *B*, *C* என்ற மூன்று அதிகாரிகள் தனித் தனியே அ, ஆ, இ, ஈ, உ, ஊ என்ற 6 நபர்களைப் பேட்டி கண்டு அவர்களின் தரங்களைக் கீழ்க்கண்ட விதத்தில் நிர்ணயித்துள்ளனர் என்று கொள்வோம்.

		நபர்கள்					
		அ	ஆ	இ	ஈ	உ	ஊ
அதிகாரிகள்	*A*	2	6	3	1	4	5
	B	5	1	6	4	3	2
	C	4	6	2	5	1	3
	R_j	11	13	11	10	8	10

இங்கு $N = 6,\ k = 3,\ \sum_{j=1}^{N=6} = 11 + 13 + 11 + 10 + 8 + 10$

$= 63$ ஆகும்.

$$\frac{\sum_{j=1}^{N} R_j}{N} = \frac{\sum_{j=1}^{6} R_j}{6} = \frac{63}{6} = 10 \cdot 5$$

எனவே, $s = \Sigma \left\{ R_j - \frac{\Sigma R_j}{N} \right\}^2$

$$\{ (11-10\cdot5)^2 + (13-10\cdot5)^2 + (11-10\cdot5)^2 + (10-10\cdot5)^2 + (8-10\cdot5)^2 + (10-10\cdot5)^2 \}$$

$$= \{0\cdot25 + 6\cdot25 + 0\cdot25 + 0\cdot25 + 6\cdot25 + 0\cdot25\}$$

$$= 13\cdot50$$

இவ்வாறாக, $W = \left\{ \frac{s}{\frac{1}{12} k^2 (N^3 - N)} \right\}$

$$= \frac{13\cdot50}{\frac{1}{12} (3)^2 (6^3 - 6)}$$

$$= \frac{162}{9\,(216 - 6)}$$

$$= \frac{18}{210} = \frac{6}{70} = 0{\cdot}085$$

எனவே, $r_{s\,\text{சராசரி}} = \dfrac{kW - 1}{k - 1}$

$$= \frac{3\,(0{\cdot}085) - 1}{3 - 1} = \frac{0{\cdot}255 - 1}{2}$$

$$= \frac{-\,0{\cdot}745}{2} = -\,0{\cdot}37$$ ஆகும்.

முதலில் கூறப்பட்ட முறையின் மூலமாகும், $r_{s\,\text{சராசரி}}$-ஐக் காணலாம். இதற்கு A, B; B, C; C, A ஆகியவைகளுக்கிடையே யுள்ள ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழுவைக் காணவேண்டும். அதாவது, r_{SAB}, r_{SBC}, r_{SCA} ஆகியவைகளைக் காணவேண்டும். இப்பொழுது r_{SAB} ஐக் காணக் கீழ்க்கண்ட அட்டவணையை அமைப்போம்.

நபர்கள்	தரங்கள்			
	A	B	d	d^3
அ	2	5	−3	9
ஆ	6	1	5	25
இ	3	6	−3	9
ஈ	1	4	−3	9
உ	4	3	1	1
ஊ	5	2	3	9
				$\Sigma d^2 = 62$

$$\text{எனவே, } r_{SAB} = 1 - \left\{\frac{6\ \Sigma d^2}{N^3 - N}\right\}$$

$$= 1 - \left\{\frac{6\ (62)}{6^3 - 6}\right\}$$

$$= 1 - \frac{62}{35}$$

$$= 1 - 1{\cdot}77$$

$$= \quad -0{\cdot}77 \text{ ஆகும்.}$$

பிறகு, இதைப்போன்று r_{SBC} ஐக் காணக் கீழ்க்கண்ட அட்டவணையை அமைப்போம்.

நபர்கள்	தரங்கள்			
	B	C	d	d^2
அ	5	4	1	1
ஆ	1	6	−5	25
இ	6	2	4	16
ஈ	4	5	−1	1
உ	3	1	2	4
ஊ	2	3	−1	1
				$\Sigma d^2 = 48$

$$\text{எனவே, } r_{SBC} = 1 - \left\{\frac{6\ \Sigma d^2}{N^3 - N}\right\}$$

$$= 1 - \left\{\frac{6\ (48)}{6^3 - 6}\right\}$$

$$= 1 - \frac{48}{35}$$

$$= 1 - 1{\cdot}37 = -\ 0{\cdot}37 \text{ ஆகும்.}$$

இறுதியாக r_{SCA} ஐக் காணக் கீழ்க்கண்ட அட்டவணையை அமைப்போம்.

நபர்கள்	தரங்கள்			
	C	A	d	d^2
அ	4	2	2	4
ஆ	6	6	0	0
இ	2	3	− 1	1
ஈ	5	1	4	16
உ	1	4	− 3	9
ஊ	3	5	− 2	4
				$\Sigma d^2 = 34$

எனவே, $r_{SCA} = 1 - \left\{\frac{6\ \Sigma d^2}{N^3 - N}\right\}$

$$= 1 - \left\{\frac{6 \times 34}{6^3 - 6}\right\}$$

$$= 1 - \frac{34}{35}$$

$$= 1 - 0{\cdot}97 = 0{\cdot}03$$

இப்பொழுது, $r_{\text{சராசரி}}$

$$= \frac{r_{SAB} + r_{SBC} + r_{SCA}}{k}$$

$$= \frac{(-0{\cdot}77) + (-0{\cdot}37) + (0{\cdot}03)}{3}$$

$$= \frac{-1{\cdot}11}{3} = -0{\cdot}37$$

இவ்வாறு இரு முறைகளிலும் $r_{\text{சராசரி}}$ மதிப்பைப் பெறலாம்.

குறிப்பு : கண்டறிந்த மதிப்புகள் ஒத்தவைகளாக இருந்தால் அவைகளுக்குச் சராசரியான முறையில் தரங்கள் பகிர்ந்தளிக்கப் பட்டு W-ஐக் கணக்கிடும்போது ஒத்தவைகளுக்கான திருத்தத்தைச் செய்கிறோம்.

$$\text{திருத்தப்பட்ட } W = \left[\frac{s}{\left\{\frac{1}{12}k^2\ (N^3 - N) - k\sum_{T} T\right\}}\right]$$

$$\text{இதில் } T = \Sigma\,\frac{(t^3 - t)}{12}\ ;$$

t என்பது ஒரே தரங்களைப் பெற்ற மதிப்புகளின் எண்ணிக்கை யாகும்.

$\sum_{T} T$ என்பது k மாறிகளின் T மதிப்புகளின் மொத்தமாகும்.

உதாரணமாக X, Y, Z ஆகிய மாறிகளின் T மதிப்புகளின் மொத்தமான $\sum_{T} T = T_X + T_Y + T_Z$ ஆகும்.

மாதிரி : கையெழுத்துப் போட்டியில் 10 மாணவர்களுக்கு X, Y, Z என்ற நீதிபதிகள் அளித்த தரங்கள் பின்வருமாறு :

மாறி	நபர்கள்									
	I	II	III	IV	V	VI	VII	VIII	IX	X
X	3	2	7·5	4·5	6	7·5	9	4.5	10	1
Y	4·5	2·5	8	4·5	9	10	6·5	1	6·5	2·5
Z	4·5	4·5	4·5	4·5	8	8	8	1	10	2

மேலுள்ள விவரங்களைக் கொண்டு கெண்டாலின் இணக்கத் தொடர்புக்கான W-ஐக் காண்போம்.

கொடுக்கப்பட்ட விவரங்களிலிருந்து R_j-ஐப் பின்வருமாறு அட்டவணைப்படுத்துவோம்.

மாறி	நபர்கள்									
	I	II	III	IV	V	VI	VII	VIII	IX	X
X	3	2	7·5	4·5	6	7·5	9	4·5	10	1
Y	4·5	2·5	8	4·5	9	10	6·5	1	6·5	2·5
Z	4·5	4·5	4·5	4·5	8	8	8	1	10	2
R_j	12	9	20	13·5	23	25·5	23·5	6·5	26·5	5·5

இப்பொழுது,

$$\frac{\Sigma R_j}{N} = \frac{\{12 + 9 + 20 + 13\cdot5 + 23 + 25\cdot5 + 23\cdot5 + 6\cdot5 + 26\cdot5 + 5\cdot5\}}{10}$$

$$= \frac{165\cdot0}{10} = 16\cdot5$$

எனவே, $s = \{(12 - 16\cdot5)^2 + (9 - 16\cdot5)^2 + (20 - 16\cdot5)^2 + (13\cdot5 - 16\cdot5)^2 + (23 - 16\cdot5)^2 + (25\cdot5 - 16\cdot5)^2 + (23\cdot5 - 16\cdot5)^2 + (6\cdot5 - 16\cdot5)^2 + (26\cdot5 - 16\cdot5)^2 + (5\cdot5 - 16\cdot5)^2\}$

$= 591\cdot00$ ஆகும்.

X, Y, Z ஆகிய மாறிகளின் தரங்களில் ஒத்தவை உள்ளன என்பதால் திருத்தத்தைப் பெற்ற *W*-ஐக் காணவேண்டும்.

முதலில், $T_X = \frac{\Sigma(t_x^3 - t_x)}{12}$

$$= \frac{(2^3 - 2)}{12} + \frac{(2^3 - 2)}{12}$$

$$= \frac{1}{2} + \frac{1}{2}$$

$$= 1$$

$$T_Y = \frac{\Sigma\,(t_y{}^3 - t_y)}{12}$$

$$= \frac{(2^3 - 2)}{12} + \frac{(2^3 - 2)}{12} + \frac{(2^3 - 2)}{12}$$

$$= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = 1 \cdot 5$$

$$T_Z = \frac{\Sigma\,(t_z{}^3 - t_z)}{12}$$

$$= \frac{(4^3 - 4)}{12} = \frac{(3^3 - 3)}{12}$$

$$= \frac{60}{12} + \frac{24}{12}$$

$$= 5 + 2 = 7$$

எனவே, $\sum_T T = T_X + T_Y + T_Z$

$$= 1 + 1 \cdot 5 + 7$$

$= 9 \cdot 5$ ஆகும்.

இவ்வாறாகப் பெறப்பட்ட $\sum_T T = 9 \cdot 5$ஐப் பயன்படுத்தி திருத்தப்பட்ட W-ஐ,

$$W = \left[\frac{s}{\frac{1}{12} k^2 (N^3 - N) - k \sum_T T}\right]$$

$$= \left[\frac{591}{\{\frac{1}{12} (3)^2 (10^3 - 10) - 3\,(9 \cdot 5)\}}\right]$$

$$= \frac{591}{742 \cdot 5 - 28 \cdot 5}$$

$= \frac{591}{714} = 0 \cdot 828$ ஆகும்.

கெண்டாலின் இணக்கக் கெழுவிற்கு மிகைத்தன்மைச் சோதனை (Testing the Significance of Kendall's Coefficient of Concordance)

சிறிய கூறுகளுக்குக் கண்டறிந்த W-ன் மதிப்பு மிகைத்தன்மை செய்கிறதா என்பதை, H_0-ன் கீழ் s-ன் மதிப்பு அதிகமாக இருக்கும் படி தொடர்புடைய நிகழ்தகவின் மதிப்பை அறிவதன் மூலம் காணலாம். k மாறிகளுக்கு, N தரங்களின் முடியுமான எல்லா வழிகளுக்கும் s-ன் மாதிரிப் பரவலைக் (sampling distribution) கண்டால், $(N!)^k$ முடியுமான தர அடுக்குகளைப் பெறலாம். இவைகளைப் பயன்படுத்தி k மாறிகளின் தரங்கள் சார்பற்றன என்ற சூனிய எடுகோளைச் சோதனை செய்யலாம். H_0-ன் கீழ், s-ன் பரவலுக்கான தீர்வு கட்டமான மதிப்புகளை (critical values) அறிய முடியும். அட்டவணை R-ல், $\alpha = 0 \cdot 05$, $0 \cdot 01$ என்ற மிகைத்தன்மை மட்டத்திற்கேற்ற W-ன் மிகைத்தன்மைக்கு s-ன் மதிப்புகள் கொடுக்கப்பட்டுள்ளன.

கண்டறிந்த s-ன் மதிப்பு அட்டவணை R-ல் α என்ற மிகைத் தன்மை மட்டத்திற்கும், N-க்கும், k-க்கும் ஏற்ற மதிப்பிற்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் சூனிய எடு கோளான H_0-ஐ α என்ற மிகைத்தன்மை மட்டத்தில் மறுக் கிறோம். உதாரணமாக, $k = 3$, $N = 6$-க்கு s-ன் கண்டறிந்த மதிப்பு $13 \cdot 50$ என்று கண்டோம். இப்பொழுது, அட்டவணை R-லிருந்து $k = 3$, $N = 6$-க்கு $\alpha = 0 \cdot 05$ என்ற மிகைத் தன்மைக்கு ஏற்ற மதிப்பு $103 \cdot 9$ ஆகும். கண்டறிந்த s-ன் மதிப் பான $13 \cdot 50$ ஆனது அட்டவணையின் மதிப்பான $103 \cdot 9$-ஐ விடக் குறைவாக உள்ளதால் மூன்று மாறிகளின் தரங்கள் சார்பற்றன வாக உள்ளன என்ற சூனிய எடுகோளை ஏற்றுக்கொள்கிறோம்.

$N > 7$ எனில், அக் கூறுகளைப் பெரிய கூறுகள் (large samples) என்று கூறுகிறோம். பெரிய கூறுகளுக்கு,

$$\chi^2 = \frac{s}{\frac{1}{12} kN(N+1)}$$

என்ற புள்ளியியல் அளவையானது $(N-1)$ வரையற்ற பாகையைப் பெற்ற கைவர்க்கப் பரவலாக அமைகிறது. அட்டவணை C-ல் H_0-ன் கீழ் $\chi^2 \geqslant$ கண்டறிந்த χ^2-ன் மதிப்புக்கு ஏற்ற தீர்வுகட்டமான மதிப்புகள் கொடுக்கப்பட்டுள்ளன.

மற்றொரு விதமாக $\dfrac{s}{\frac{1}{12} k N (N + 1)} = k (N-1) W$ ஆகும்.

எனவே, $\chi^2 = k (N-1) W$ ஆகும்.

இச் சூத்திரம் மிக எளிதாக அமைந்துள்ளதால் இதனைப் பயன்படுத்தி χ^2-ன் மதிப்பைக் காணலாம். கண்டறிந்த χ^2-ன் மதிப்பானது அட்டவணை C-யிலுள்ள $(N-1)$ வரையற்ற பாகைக்கும், α என்ற மிகைத்தன்மை மட்டத்திற்கும் ஏற்ற மதிப்புக்குச் சமமாகவோ அல்லது அதிகமாகவோ இருந்தால் k மாறிகளின் தரங்கள் சார்பற்றனவாக அமைந்துள்ளன என்ற சூனிய எடுகோளை α என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

மாதிரி : கையெழுத்துப் போட்டியில் 10 மாணவர்களுக்கு X, Y, Z என்ற நீதிபதிகள் அளித்த தரங்கள் பின்வருமாறு :

மாறி	நபர்கள்									
	I	II	III	IV	V	VI	VII	VIII	IX	X
X	3	2	7·5	4·5	6	7·5	9	4·5	10	1
Y	4·5	2·5	8	4·5	9	10	6·5	1	6·5	2·5
Z	4·5	4·5	4·5	4·5	8	8	8	1	10	2

மேலுள்ள விவரங்களிலிருந்து $k = 3$, $N = 10$-க்கு $W = 0{\cdot}828$ என்று முன்பு கண்டோம்.

இங்கு $N = 10 > 7$ என்பதால்,

$\chi^2 = k(N-1)W$ என்ற சூத்திரத்தைப் பயன்படுத்துகிறோம்.

எனவே, $\chi^2 = k(N-1)W$

$= 3(10-1)(0{\cdot}828)$

$= 22{\cdot}356$ ஆகும்.

அட்டவணை C-யிலிருந்து $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்திற்கும் $(N-1) = (10-1) = 9$ வரையற்ற பாகைக்கும் ஏற்ற χ^2-ன் மதிப்பு $16{\cdot}92$ ஆகும். கண்டறிந்த χ^2-ன் மதிப்பான $22{\cdot}356$ ஆனது அட்டவணை மதிப்பான $16{\cdot}92$-ஐ விட அதிகமாக உள்ளதால், X, Y, Z என்ற மூன்று மாறிகளின் தரங்கள் சார்

பற்றன என்ற சூனிய எடுகோளை $\alpha = 0{\cdot}05$ என்ற மிகைத்தன்மை மட்டத்தில் மறுக்கிறோம்.

பயிற்சி 6

1. கீழ்க்கண்ட பட்டியலிலிருந்து பெற்றோர்களின் உயரத்திற்கும் அவர்களுடைய பிள்ளைகளின் உயரத்திற்கும் உள்ள உறவினை இணைப்புக் கெழுவுக்கான சூத்திரத்தைப் பயன்படுத்திக் காண்க. மேலும் அதற்கான மிகைத்தன்மைச் சோதனையும் செய்க.

		பெற்றோர்கள்			
		உயரமான	நடுத்தரமான	குட்டையான	மிகவும் குட்டையான
பிள்ளைகள்	உயரமான	25	30	20	2
	நடுத்தரமான	19	125	85	12
	குட்டையான	3	140	165	125
	மிகக் குட்டையான	3	37	68	151

2. ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழு என்றால் என்ன ? கண்டறிந்தவை ஒத்தவைகளாக இருந்தால் அதற்கான திருத்தத்தைப் பெற்ற ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழுவைக் காணும் முறையை விளக்குக.

3. ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழுவைப் பயன்படுத்தி அதற்கான மிகைத்தன்மைச் சோதனையைச் செய்யும் முறையினை விளக்குக.

4. 10 மாணவர்கள் *X*, *Y* என்ற இரு உளவியல் சோதனைகளில் பெற்ற மதிப்பெண்கள் கீழ்வருமாறு :

மாணவர்கள்	சோதனை *X*	சோதனை *Y*
a	100	170
b	88	193
c	108	178
d	94	183
e	102	166
f	114	164
g	109	148
h	85	187
i	84	166
j	87	128

இவ் விவரங்களைக் கொண்டு ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழுவைக் காண்க. மேலும், அக் கெழுவிற்கான மிகைத்தன்மைச் சோதனையைச் செய்க.

5. தையல் நூல்களைத் தயாரிக்கும் நிறுவனம் 12 வகையான நூல்களைத் தயாரிக்கிறது. இந் நூல்களின் தரத்திற்கு ஏற்ப *X*, *Y* என்ற இரு பரிசோதகர்கள் கீழ்க்கண்டவாறு மதிப்புகள் கொடுத்தனர்:

நூல்களின் வகை	பரிசோதகர் *X*	பரிசோதகர் *Y*
A	20	62
B	20	66
C	21	59
D	21	57
E	23	85
F	24	108
G	25	106
H	26	76
I	27	82
J	28	112
K	28	74
L	32	101

மேலுள்ள விவரங்களைப் பயன்படுத்தி *X*, *Y* ஆகிய மாறிகளுக்கிடையேயுள்ள, ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழுவைக் காண்க. மேலும், இம் மாறிகளுக்கிடையே முழுமைத் தொகுதி

யில் தொடர்பு இருக்கிறதா என்பதை இக் கெழுவிற்கான மிகைத் தன்மைச் சோதனை செய்து காண்க.

6. கெண்டாலின் தர ஒட்டுறவுக் கெழு என்றால் என்ன ? இதற்கும் ஸ்பியர்மேனின் தர ஒட்டுறவுக் கெழுவிற்கும் உள்ள ஒப்புமை என்ன ?

7. *X*, *Y* என்ற இரு நீதிபதிகள் 10 மாணவர்களுக்கு ஒரு போட்டியில் அளித்த மதிப்பெண்கள் பின்வருமாறு :

X :	14,	14,	17,	17,	17,	19,	26,	27,	31,	35
Y :	18,	26,	18,	18,	18,	26,	30,	22,	35,	30

மேலுள்ள விவரங்களைப் பயன்படுத்திக் கெண்டாலின் தர ஒட்டுறவுக் கெழுவைக் கண்டு அக் கெழுவிற்கான மிகைத்தன்மைச் சோதனையை மேற்கொள்க.

8. ஒரு வகுப்பில் படிக்கும் 12 மாணவர்கள் கணக்கு, புள்ளியியல், பொருளாதாரம் ஆகிய தேர்வுகளில் பெற்ற மதிப்பெண்கள் கீழ்வருமாறு :

மாணவர்கள்	மதிப்பெண்கள்		
	கணக்கு	புள்ளியியல்	பொருளாதாரம்
அ	27	18	19
ஆ	36	54	19
இ	18	45	38
ஈ	9	9	46
உ	72	90	54
ஊ	99	81	78
எ	90	72	27
ஏ	54	27	81
ஐ	63	36	98
ஒ	100	100	85
ஓ	45	63	88
ஔ	81	99	64

மேலுள்ள விவரங்களிலிருந்து $r_{xy \cdot z}$ என்ற கெண்டாலின் பகுதி தர ஒட்டுறவுக் கெழுவைக் காண்க. அதற்கான மிகைத் தன்மைச் சோதனையை மேற்கொள்க.

9. A, B, C என்ற மாறிகளுக்கான 12 தரங்கள் கீழே கொடுக்கப்பட்டுள்ளன:

தரங்கள்		
A	B	C
3	2	1·5
4	6	1·5
2	5	3·5
1	1	3·5
8	10	5·0
11	9	6·0
10	8	7·0
6	3	8·0
7	4	9·0
12	12	10·5
5	7	10·5
9	11	12·0

மேலுள்ள விவரங்களைப் பயன்படுத்தி r_{ABC} என்ற கெண்டாலின் பகுதி ஒட்டுறவுக் கெழுவைக் காண்க. பின்பு அக் கெழுவிற்கான மிகைத்தன்மைச் சோதனையைச் செய்க.

10. 6 மனுதாரர்களைப் பேட்டி கண்ட X, Y, Z என்ற அதிகாரிகள் அளித்த தரங்கள் கீழ்வருமாறு :

		மனுதாரர்கள்					
		a	b	c	d	e	f
அதிகாரிகள்	X	1	6	3	2	5	4
	Y	1	5	6	4	2	3
	Z	6	3	2	5	4	1

மேலுள்ள விவரங்களைப் பயன்படுத்திக் கெண்டாலின் இணக்கத் தொடர்புக் கெழுவையும், $r_{S\ சராசரி}$-ஐயும் கண்டு W-க்கான மிகைத்தன்மைச் சோதனையைச் செய்க.

11. பெரிய கூறுகளுக்குக் கெண்டாலின் இணக்கத் தொடர்புக் கெழுவிற்கான மிகைத்தன்மைச் சோதனை செய்யும் முறையைப் பற்றி விவரிக்க.

12. நர்ஸரி பள்ளியில் பயிலும் 20 குழந்தைகளின் பேச்சுத் திறமைக்கான தரங்களை 13 ஆசிரியைகள் தனித்தனியாக அளித்தனர். அவைகளின் விவரம் பின்வருமாறு:

ஆசிரியைகள்	குழந்தைகள்																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
a	11	2	3	4	5	6	7	8	9	10	1	12	13	14	15	16	17	18	19	20
b	9	1	16	8	5	2	6	10	4	3	11	13	7	12	17	18	19	15	14	20
c	13	2	7	5	14	9	15	16	6	11	8	10	1	4	19	12	20	3	17	18
d	18	13	10	11	4	2	5	3	9	1	14	7	6	15	16	12	19	17	8	20
e	6	14	15	8	1	4	2	9	7	10	11	5	3	16	11	13	18	17	12	19
f	10	17	5	13	15	11	7	4	9	2	18	3	6	1	19	12	16	8	14	20
g	14	9	12	6	7	2	3	10	5	4	17	8	1	15	13	16	18	11	20	19
h	16	2	13	10	7	3	4	14	6	5	17	9	1	12	8	11	20	15	18	19
i	19	2	15	6	5	7	8	10	9	3	12	4	1	13	11	14	9	18	16	17
j	12	4	16	3	10	6	14	17	15	7	19	9	1	8	5	13	11	18	2	20
k	17	14	12	8	7	2	5	10	3	4	13	9	1	18	6	15	19	16	11	20
l	15	1	13	3	5	2	14	9	6	10	8	11	19	4	7	12	18	17	16	20
m	15	3	13	2	8	1	9	12	4	6	14	10	11	7	5	18	16	17	19	20

மேலுள்ள விவரங்களைப் பயன்படுத்திக் கெண்டாலின் இணக்கத் தொடர்புக் கெழு *W*-ஐக் கண்ட பின்பு அதற்கான மிகைத்தன்மைச் சோதனை செய்க.

பிற்சேர்க்கை

(Appendix)

புள்ளியியல் அட்டவணைகள்

(Statistical Tables)

அட்டவணை A: 'இயல்நிலைப் பரவலில், z-ன் கண்டறிந்த மதிப்பைவிடப் புறக்கோடியாக அமைந்துள்ள மதிப்புகளோடு தொடர்புடைய நிகழ்தகவு மதிப்புகள்.' இதற்கான அட்டவணை, Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கம் 247-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை B: 't-ன் தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை, Fisher and Yates (6th Edition—Longman, 1974), 'Statistical Tables for Biological, Agricultural and Medical Research' என்ற புத்தகத்தின், 46ஆம் பக்கத்தில் அட்டவணை III-ன் கீழ் கொடுக்கப்பட்டுள்ளது.

அட்டவணை C: 'கைவர்க்க தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை, Fisher and Yates (6th Edition—Longman, 1974), 'Statistical Tables for Biological, Agricultural and Medical Research' என்ற புத்தகத்தின் 47ஆம் பக்கத்தில் அட்டவணை IV-ன் கீழ் கொடுக்கப்பட்டுள்ளது; அல்லது Lindley, D. V. and Miller, J. C. P. (1953), 'Cambridge Elementary Statistical Tables.' Cambridge University Press, என்ற புத்தகத்தில் கொடுக்கப்பட்டுள்ளது.

அட்டவணை D: 'ஈருறுப்புப் பரவல்' $P(Y \leq y) = \sum_{i=0}^{y} \binom{n}{i} p^i (1-p)^{n-i}$. இதற்கான அட்டவணை, William Jay Conover, 'Practical Nonparametric Statistics' என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கம் 368-லிருந்து 379 வரை கொடுக்கப்பட்டுள்ளது.

அட்டவணை E: 'கோல்மோகோராவ் சோதனைப் புள்ளியியல் அளவையின் பகுதிமானங்கள்.' இதற்கான அட்டவணை, Miller, L. H. (1956), Table of Percentage Points of Kolmogorov Statistics, Journal of the American Statistical Association என்ற பத்திரிகையில் Vol. 51, பக்கங்கள் 111—121-ல் கொடுக்கப்பட்டுள்ளது. அல்லது William Jay Conover, 'Practical Nonparametric Statistics' என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கம் 397-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை F: 'ஓட்டச் சோதனைக்கு r-ன் தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை Swed, Frieda, S., and Eisenhart, C., 1943, Tables for testing randomness of grouping in a sequence of alternatives, The Annals of Mathematical Statistics என்ற பத்திரிகையில் Vol. 14, பக்கங்கள் 83—86-ல் கொடுக்கப்பட்டுள்ளது. அல்லது, Sidney Siegel 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கங்கள் 252—253-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை G: 'வில்காக்ஸன் ஜோடிப் பொருத்தக் குறியிடு தரச் சோதனைக்கு, T-ன் தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை, Sidney Siegel 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 254-ல் கொடுக்கப்பட்டுள்ளது; அல்லது R. C. Campbell, 'Statistics for the Biologists', Cambridge University Press என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 223-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை H: 'வால்ஷ் சோதனைக்குத் தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை, Walsh, J. E. (1949), Applications of some significance tests for the median, which are valid under very general conditions, Journal of the American Statistical Association என்ற பத்திரிகையில் Vol. 44, பக்கம் 343-ல் கொடுக்கப்பட்டுள்ளது. அல்லது, Sidney Siegel 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 255-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை I: 'ஃபிஷரின் சோதனையில் D-ன் (அல்லது C-ன்) தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை Finney, D. J., 1948, The Fisher and Yates 'Test of Significance in 2×2

Contingency Tables', Biometrika என்ற பத்திரிகையில் Vol. 35, பக்கங்கள் *149—154*-ல் கொடுக்கப்பட்டுள்ளது. அல்லது Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கங்கள் *257—270*-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை J: 'மேன்-விட்னி *U*' சோதனையில், கண்டறிந்த *U*-ன் மதிப்பைவிடக் குறைவாக இருப்பதற்கான நிகழ்தகவு மதிப்புகள். இதற்கான அட்டவணை, Mann, H. B. and Whitney, D. R., 1947, On a test of whether one of 2 random variables is stochastically larger than the other, The Annals of Mathematical Statistics என்ற பத்திரிகையில் Vol. *18*, பக்கங்கள் *52—54*-ல் கொடுக்கப்பட்டுள்ளது. அல்லது Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் *271—273*-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை K: 'மேன்-விட்னி *U*' சோதனையில், *U*-ன் தீர்வு கட்டமான மதிப்புகள். இதற்கான அட்டவணை, Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கங்கள் *274—277*-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை L: '*n* சம அளவுகளைப் பெற்ற இரு கூறுகளுக்கு ஸ்மிர்னுவ் சோதனைப் புள்ளியியல் அளவையின் பகுதிமானங்கள்.' இதற்கான அட்டவணை, Birnbaum, Z. W., and Hall, R. A. (1960) Small Sample Distribution for Multi Sample Statistics of the Smirnov Type, The Annals of Mathematical Statistics என்ற பத்திரிகையில் Vol. *31*, பக்கங்கள் *710—720*-ல் கொடுக்கப்பட்டுள்ளது. அல்லது William Jay Conover, 'Practical Nonparametric Statistics', Wiley Inter-Science Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கம் *399*-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை M: '*n*, *m* ஆகிய வெவ்வேறு அளவுகளைப் பெற்ற இரு கூறுகளுக்கு ஸ்மிர்னுவின் சோதனைப் புள்ளியியலின் பகுதி மானங்கள்.' இதற்கான அட்டவணை, William Jay Conover, 'Practical Nonparametric Statistics', Wiley Inter-Science Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில் பக்கங்கள் *400*, *401* ஆகியவைகளில் கொடுக்கப்பட்டுள்ளது.

அட்டவணை N: 'ஃபிரிட்மேன் இருவழி பரவற்படி ஆய்வு தரச் சோதனையில், கண்டறிந்த χ_r^2-ன் மதிப்பைவிட அதிகமாக

இருக்கும் மதிப்புகளுக்குத் தொடர்புடைய நிகழ்தகவு மதிப்புகள்.' இதற்கான அட்டவணை, Friedman, M., 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance, Journal of the American Statistical Association என்ற பத்திரிகையில் Vol. 32, பக்கங்கள் 688—689-ல் கொடுக்கப்பட்டுள்ளது. அல்லது Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கங்கள் 280—281-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை O: 'க்ரஸ்கல்-வாலீஸின் ஒருவழி பரவற்படி ஆய்வு தரச் சோதனையில் கண்டறிந்த H-ன் மதிப்பைவிட அதிகமாக இருக்கும் மதிப்புகளுக்குத் தொடர்புடைய நிகழ்தகவு மதிப்புகள்'. இதற்கான அட்டவணை, R. C. Campbell, 'Statistics for Biologists', Cambridge University Press என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கங்கள் 220—221-ல் கொடுக்கப்பட்டுள்ளது. அல்லது Sidney Siegel, 1956 'Nonparametric Statistics for the Behavioral Sciences' என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கங்கள் 282—283-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை P: 'ஸ்பியர்மேன்' தர ஒட்டுறவுக் கெழுவான r_s-ன் தீர்வுகட்டமான மதிப்புகள். இதற்கான அட்டவணை, Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 284-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை Q: 'கெண்டாலின் தர ஒட்டுறவுக் கெழுவான S-ன் கண்டறிந்த மதிப்பைவிட அதிகமாக இருக்கும் S-ன் மதிப்புகளுக்குத் தொடர்புடைய நிகழ்தகவு மதிப்புகள்.' இதற்கான அட்டவணை, Sidney Siegel, 1957, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 285-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை R: 's என்ற கெண்டாலின் இணக்கக் கெழுவின் தீர்வுகட்டமான மதிப்புகள்.' இதற்கான அட்டவணை, Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw-Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 286-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை S: 'காரணியப் பெருக்கங்கள்.' இதற்கான அட்டவணை, Sidney Siegel, 1956, 'Nonparametric Statistics for the Behavioral Sciences', McGraw–Hill Book Co., Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 287-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை T: 'திட்ட இயல்நிலை ராண்டம் மாறி W-ன் பகுதி மானங்கள்: W_p-க்கான $P(W \leqslant W_p) = p$ அல்லது $P(W > W_p) = 1-p$ இயல்நிலைப் பரவல்.' இதற்கான அட்டவணை, Pearson, E. S. and Hartley, H.O. (1962), 'Biometrika Tables for Statisticians', Vol. 1, 2nd Edition, Cambridge, England என்ற புத்தகத்தில் கொடுக்கப்பட்டுள்ளது. அல்லது W. J. Conover, 'Practical Nonparametric Statistics', Wiley Inter-Science, Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 366-ல் கொடுக்கப்பட்டுள்ளது.

அட்டவணை U: 'வில்லிஃபார்ஸ் சோதனைப் புள்ளியியல் அளவையின் பகுதிமான மதிப்புகள்.' இதற்கான அட்டவணை, William Jay Conover 'Practical Nonparametric Statistics', Wiley Inter-Science, Inc. என்ற புத்தகத்தின் பிற்சேர்க்கையில், பக்கம் 398-ல் கொடுக்கப்பட்டுள்ளது.

மேற்கோள் நூற்பட்டியல்

நூல்கள்
(Books)

1. Conover, William Jay (1971), 'Practical Nonparametric Statistics.' John Wiley & Sons Inc., New York.

2. Fraser, D. A. S. (1957), 'Nonparametric Methods in Statistics', John Wiley, New York.

3. Freund, J. E. (1962). 'Mathematical Statistics', Prentice-Hall, Englewood Cliffs, N.J.

4. Jaroslav Hajek (1969), 'A Course in Nonparametric Statistics', Holden-Day, San Francisco.

5. Kendall, M.G. (1955), 'Rank Correlation Methods', 2nd ed., Hafner, New York.

6. Noether, G. E. (1967a), 'Elements of Nonparametric Statistics', John Wiley, New York.

7. Siegel, S. (1956), 'Nonparametric Statistics for the Behavioral Sciences', 'McGraw-Hill, New York.

பத்திரிகைகள்
(Journals)

1. The Annals of Mathematical Statistics: Vol. 9, 1938. Distributions of sums of squares of rank differences for small numbers of individuals—Olds, E. G., pp. 133—148.

2. The Annals of Mathematical Statistics: Vol. 11, 1940. A comparison of alternative tests of significance for the problem of *m* rankings—Friedman, M., pp. 86—92.

3. The Annals of Mathematical Statistics: Vol. 14, 1943. Tables for testing randomness of grouping in a sequence of alternatives—Swed, Frieda, S. and Eisenhart, C., pp. 83—86.

4. The Annals of Mathematical Statistics: Vol. 18, 1947. On a test of whether one of two random variables is stochastically larger than the other—Mann, H. B. and Whitney D. R., pp. 52—54.

5. The Annals of Mathematical Statistics: Vol. 20, 1949. The 5% significance levels for sums of squares of rank differences and a correction—Olds, E.G., pp. 117—118.

6. The Annals of Mathematical Statistics: Vol. 23, 1952. Distribution table for the deviation between two sample cumulatives. Massey, F. J., pp. 435—441.

7. The Annals of Mathematical Statistics: Vol. 31, 1960. Small sample distributions for multiple sample statistics of the Smirnov type – Birnbaum, Z. W. and Hall, R. A. pp. 710—720.

8. Journal of the American Statistical Association; Vol. 32, 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance—Friedman, M., pp. 688—689.

9. Journal of the American Statistical Association: Vol. 44, 1949. Applications of some significance tests for the median which are valid under very general conditions—Walsh, J. E., p. 343.

10. Journal of the American Statistical Association: Vol 47, 1952. Use of ranks in one criterion variance analysis—Kruskal, W. H. and Wallis, W. A., pp. 614—617.

11. Journal of the American Statistical Association: Vol. 51, 1956. Table of Percentage Points of Kolmogorov Statistics—Miller, L. H., pp. 111—121.

12. Journal of the American Statistical Association: Vol. 62, 1967, On the Kolmogorov-Smirnov test for normality with mean and variance unknown—Lilliefors, H. W., pp. 399—402.

13. Biometrika: Vol. 35, 1948. The Fisher-Yates test of significance in 2×2 contingency tables—Finney, D. J., pp. 149—154.

14. Bulletin of the Institute of Educational Research at Indiana University: Vol. 1, No. 2, 1953. 'Extended tables for the Mann—Whitney Statistic', Auble, D.

15. Journal of the American Statistical Association: Vol. 57, 1962. A distribution - free test of independence with a sample of paired observations - Gerald J. Glasser, pp. 116—133.

கலைச்சொற்கள்

A

Analysis of variance — பரவற்படி ஆய்வு
Arithmetic mean — சராசரி
Ascending order — ஏறு வரிசை
Assumption — தற்கோள்
Attribute — பண்பு

B

Binomial distribution — ஈருறுப்புப் பரவல்
Bivariate frequency distribution — இருமாறி அலைவுப் பரவல்

C

Central tendency — மையப்போக்கு
Chi square distribution — கைவர்க்கப் பரவல்
Class — பிரிவு
Class interval — பிரிவு இடைவெளி
Classification — பாகுபாடு
Classification one-way — ஒருவழிப் பாகுபாடு
Coefficient — கெழு
Coefficient of contingency — இணைப்புக் கெழு
Coefficient of skewness — கோட்ட அளவைக் கெழு
Coefficient of variation — மாறுபாட்டுக் கெழு
Confidence limit — நம்பிக்கை எல்லை
,, lower — கீழ்நம்பிக்கை எல்லை
,, upper — மேல் நம்பிக்கை எல்லை
Contingency table — இணைப்பட்டியல்
Correlation — ஒட்டுறவு
Correlation coefficient — ஒட்டுறவுக் கெழு
,, multiple — பல்தர ஒட்டுறவு
,, negative — எதிர்மறை ஒட்டுறவு
,, partial — பகுதி ஒட்டுறவு
,, positive — நேரிடை ஒட்டுறவு
,, rank — தர ஒட்டுறவு
,, table — ஒட்டுறவுப் பட்டியல்
Critical region — தீர்வுகட்டமான வெளி
Critical value — தீர்வுகட்டமான மதிப்பு
Cumulative — குவிவு

Cumulative curve	—	குவிவு வளைகோடு
,, distribution	—	குவிவுப் பரவல்
,, frequency	—	குவிவு அலைவெண்
,, frequency curve	—	குவிவு அலைவெண் வளைகோடு
Curve	—	வளைகோடு

D

Data	—	விவரங்கள்
Degree of freedom	—	வரையற்ற பாகை
Dependent event	—	சார்புடைய நிகழ்ச்சி
Descending order	—	இறங்கு வரிசை
Difference	—	வேறுபாடு
Distribution	—	பரவல்
,, binomial	—	ஈருறுப்புப் பரவல்
,, frequency	—	அலைவெண் பரவல்
,, normal	—	இயல்நிலைப் பரவல்

E

Emperical distribution	—	அனுபவப் பரவல்
Estimated value	—	மதிப்பளவை
Event	—	நிகழ்ச்சி
Event dependent	—	சார்புடைய நிகழ்ச்சி
,, favourable	—	சாதகமான நிகழ்ச்சி
,, independent	—	சார்பற்ற நிகழ்ச்சி
Expected value	—	எதிர்பார்க்கும் மதிப்பு

F

Frequency	—	அலைவெண், அலைவு
Frequency distribution	—	அலைவெண் பரவல்
Function	—	சார்பலன்

H

Heterogeneous	—	பலபடித்தான
Homogeneous	—	ஒருபடித்தான
Hypothesis	—	கருதுகோள், எடுகோள்

I

Identically distributed	—	ஒரே மாதிரியான பரவலைப் பெற்ற
Independent	—	சார்பற்ற
Independent event	—	சார்பற்ற நிகழ்ச்சி
Inference	—	உய்த்துணர்வு

K

Kurtosis	—	தட்டை அளவை

L

Large sample — பெருங்கூறு
Level of significance — மிகைத்தன்மை மட்டம்
Lower confidence limit — கீழ் நம்பிக்கை எல்லை
Lower quartile — கீழ்க் கால்மானம்

M

Measures of central tendency — மையப்போக்கு அளவைகள்
Measures of skewness — கோட்ட அளவைகள்
Median — இடைநிலை
Minimum — நீசம்
Multiple correlation — பல்தர ஒட்டுறவு
Mutually exclusive events — ஒன்றையொன்று விலக்கும் நிகழ்ச்சிகள்

N

Nominal scale — பெயரளவிலான அளவுத் திட்டம்
Non-parametric — சுட்டுறுப்பைச் சாராத
Normal curve — இயல்நிலை வளைகோடு
Normal distribution — இயல்நிலைப் பரவல்
Null hypothesis — மறுக்கத்தகு எடுகோள், சூனிய எடுகோள்

O

Observation — கண்டறிதல்
Observed data — கண்டறிந்த விவரங்கள்
Ogive — அலைவெண் குவிவு
One-way classification — ஒருவழிப் பாகுபாடு
Ordinal scale — வரிசையளவுத் திட்டம்

P

Parameter — சுட்டுறுப்பு
Partial correlation — பகுதி ஒட்டுறவு
Personality — ஆளுமை
Population — முழுமைத்தொகுதி
Power curve — ஆக்கத்திறன் வளைகோடு
Probability — நிகழ்தகவு

Q

Quartile — கால்மானம்

R

Random sample — ராண்டம் மாதிரி
Random numbers — ராண்டம் எண்கள்
Rank — தரம்

Range — வீச்சு
Rank correlation — தர ஒட்டுறவு
Rates — வீதங்கள்
Reaction — பிரதிக் கிரியை
Related — தொடர்புடைய

S

Sample — கூறு, மாதிரி
,, average — கூறுச் சராசரி
,, random — சமவாய்ப்புக் கூறு
Sampling distribution — கூறுப் பரவல், மாதிரிப் பரவல்
Scale — அளவுத் திட்டம்
Scores — மதிப்புகள்
Shape — வடிவம்
Standard deviation — திட்ட விலக்கம்
Standard error — திட்டப் பிழை
Statistic — புள்ளியியல் அளவை
Statistical inference — புள்ளியியல் உய்த்துணர்வு
Statistics — புள்ளியியல்
Symmetrical — சமச்சீரான
Symmetrical distribution — சமச்சீரான பரவல்

T

Table — அட்டவணை, பட்டியல்
Tabulate — பட்டியலமை
Test for independence — சார்பற்ற தன்மைக்கான சோதனை
Test of significance — மிகைத்தன்மைச் சோதனை
Theoretical distribution — அறிமுறைக்குரிய பரவல்
Ties — ஒத்தவை

U

Unbiased — நடுநிலையான, சாராத
Uniform distribution — சீரான பரவல்
Unique — ஒருதனி
Unit — அலகு

V

Variate, variable — மாறி
,, dependent — சார்புடை மாறி
,, independent — சார்பற்ற மாறி
Variation — வேறுபாடு
Variance — மாறுபாடு